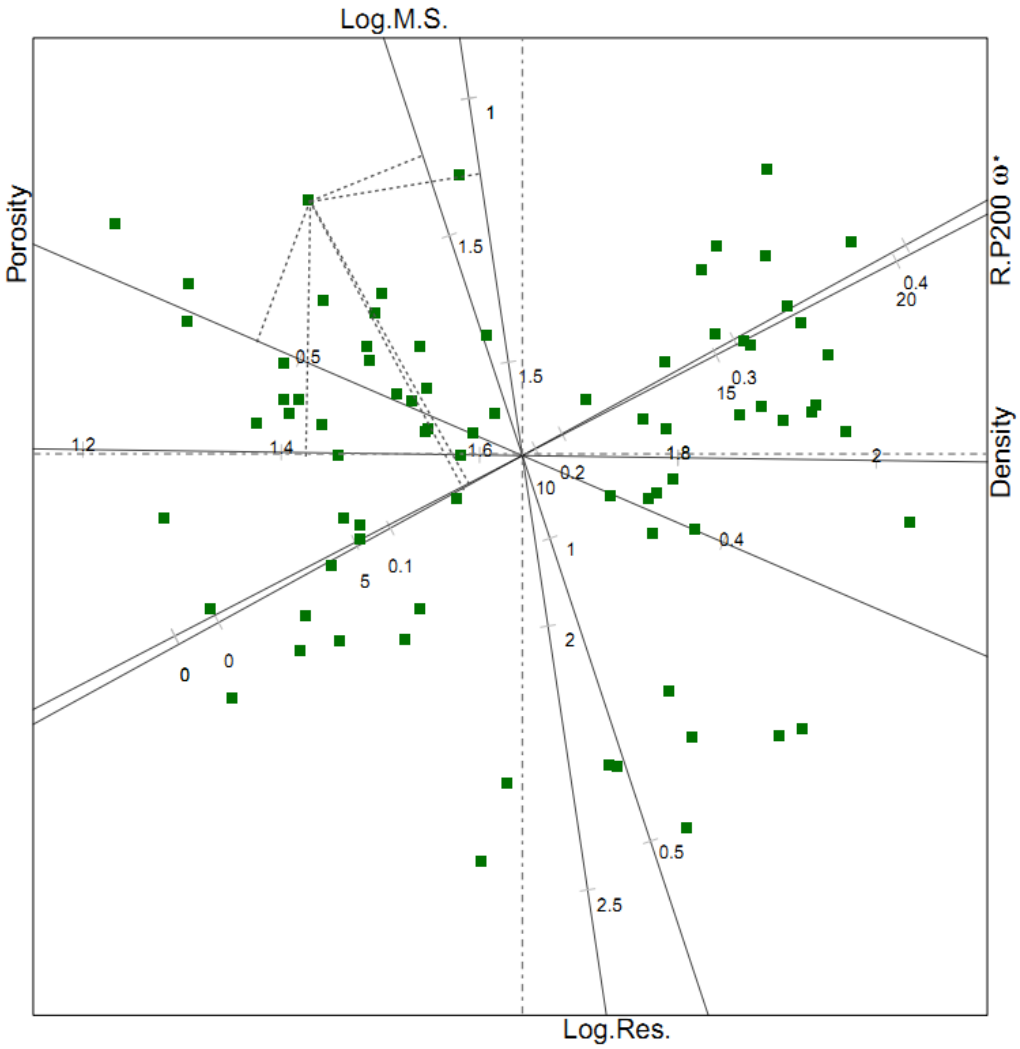
TU Clausthal

Clausthal University of Technology

An investigation on petrophysical and geotechnical properties of soils using multivariate statistics



Dissertation 2012

Pham Quy Ngoc

An investigation on petrophysical and geotechnical properties of soils using multivariate statistics

D i s s e r t a t i o n

zur Erlangung des Grades eines

Doktors der Naturwissenschaften

vorgelegt von

Phạm Quý Ngọc

aus Nam Dinh, Vietnam

genehmigt von der Fakultät für

Energie- und Wirtschaftwissenschaften

der Technischen Universität Clausthal

Tag der mündlichen Prüfung

12. July 2012

Die Arbeit wurde am Institut für Geophysik der Technischen Universität Clausthal angefertigt.

Vorsitzender der Promotionskommission:	Prof. Dr. rer. nat. H.-J. Gursky
Hauptberichterstatter:	Prof. Dr. rer. nat. habil. A. Weller
Mitberichterstatter:	Prof. Dr.-Ing. N. Meyer

Acknowledgments

The author wishes to express his deepest gratitude to his adviser, Prof. Dr. Andreas Weller, for his advises, encouragement and continuous guidance through out this research work. My grateful appreciation are also extended to Prof. Dr. Norbert Meyer for his valuable suggestions, guidance and for serving as member of the examination committee.

I am thankful to Dr. Carl-Dietrich Sattler from the Institute of Geology and Paleontology, Clausthal University of Technology for performing the X-ray diffraction investigation on soil samples and his guidance on interpretation. My sincere thanks are also extended to all colleagues in the Institute of Geophysics, Clausthal University of Technology. Great thanks to Dr. Wolfgang Debschütz, Sven Nordsik, Marcus Möller for their guidance in performance of petrophysical measurements.

Words of appreciation are extended to Dr. Tran Canh, Institute of Geological Sciences, Vietnam Academy of Science and Technology (VAST) for his assistance in collecting soil samples at the dike monitoring system in Ngo Xa, Vietnam. My thanks are also extend to Dr. Ronald Lewis, Planungsgesellschaft Scholz + Lewis mbH (PGS) for his assistance in collecting soil samples in Sachsen-Anhalt, Germany.

I also would like to thank the government of Vietnam for a partial funding in pursuing the PhD program at the Institute of Geophysics, Clausthal University of Technology.

Finally, I would express my great gratitude to my wife Le, T. Thu Hien, my daughter Pham T. Nhu Anh and my family. This dissertation could not have been finished without their endless love, encouragement and support.

Summary

The application of geophysical methods in environmental and civil engineering investigation has become more important in recent decades. A successful application of geophysical methods in investigation of subsoil requires an adequate knowledge on petrophysical and geotechnical properties of the soil and their relationships.

Various types of soil samples were collected in different locations and depths in Vietnam and Germany. The soil samples originate from the river dikes and adjacent foundation of civil engineering works. Geotechnical properties such as water content, density, Atterberg limits, clay fraction, shear strength parameters and hydraulic conductivity of soil were investigated in laboratory. The soils from Vietnam have higher clay fraction than those from Germany. The results show that an increase of water content results in a decrease of shear strength. Clay fraction presents a linear relation to liquid limit, plasticity index and specific surface area. The logarithms of hydraulic conductivity of soils indicates a linear decrease with increasing clay fraction.

Petrophysical properties such as complex resistivity, dielectric permittivity, grain density, magnetic susceptibility and specific surface area were determined in the laboratory. The volumetric water content has a strong effect on the dielectric permittivity. Magnetic susceptibility and grain density of soils from Vietnam exhibit higher values than those of soils from Germany due to higher clay content.

When dealing with a large amount of soil samples and more than three properties, a multivariate statistical analysis should be used to analyse and visualize the data and the relations among the properties in global view. Multivariate statistical methods of principal component analysis (PCA) and cluster analysis were applied to investigate the relation between geotechnical and petrophysical properties of soils. The soil samples and their properties are compiled in a data matrix, where the rows are soil samples and the their properties are compiled in the columns. Matrix manipulation

algorithms are applied to reduce the dimensionality of the problem with the least loss of information. In statistical analysis of two variables, the coefficient of determination R^2 can be considered as goodness of fit of the model. In multivariate statistics, the new concepts of overall quality, adequacy and predictivity are used to access the goodness of fit of the model.

Both row and columns of the approximation matrix can be viewed as points and axes respectively in one graph as two- or three-dimensional biplot. The soil samples are presented as points while soil properties are the axes. A specific property of a soil sample can be interpolated by a projection of sample location onto the corresponding property axis. The cosine of angle between two axes exhibit the linear relation between those two properties.

Multivariate analysis methods offer potential tools to analyse and visualize a large soil sample set with various properties. All soil samples and their properties can be visualized simultaneously in a global view. The technique also enables an effective classification of soil samples.

The multivariate statistical analysis of geotechnical data of soil samples from different locations in Vietnam has identified two groups of parameters. The correlation coefficient matrix indicates strong correlations between porosity, water content, density and friction angle. A larger porosity reduces the density and enables an increased volume of water in the pore space, whereas increased water content reduces the friction angle. These four interrelated parameters show the strongest loadings in the first principal component. The other group of parameters are cohesion and clay fraction that are characterized by a moderate correlation. A cluster analysis provides a classification into eight soil clusters. The use of only one parameter of each group seems to be a less expensive alternative of soil classification. A cross-plot of the easily determinable parameters porosity and clay fraction enables a rough differentiation between sandy, silty and clayey soils.

The petrophysical parameters can be determined by geophysical field surveys along dikes or from geophysical logging in small boreholes. The most relevant parameter that is extracted by a radar survey is the relative dielectric permittivity. The magnetic susceptibility is considered as additionally petrophysical parameter. The parameters porosity, natural raw density, and volumetric water content can be determined both from laboratory investigations or well logging. The set of six parameters was integrated for multivariate statistics. In a similar way as in the first example, density and water content dominate the first principal component. Because of the theoretically justified excellent correlation between relative dielectric permittivity and water content the real part of dielectric permittivity joins the first group of parameters. Resistivity and magnetic susceptibility, which show no correlation to the parameters of the first group, indicate the strongest loadings in the second principal component. In combination of cluster analysis and PCA biplot, soils are classified as two clusters. Using a cross-plot of density and resistivity, the soil clusters are roughly identified.

The second example has also demonstrated that the electrical resistivity is a key parameter in soil classification. This parameter is strongly related to clay content and water content of soils. Geoelectrical surveys along the crest of the dike, which are recommended for dike investigation, enable a sectioning into more sandy or clayey soils in the dike body.

List of Abbreviation and Symbols

Abbreviation

AB: An Bai, Thai Binh, Vietnam

B: Brucite sheet

CEC: Cation exchange capacity

CF: Clay fraction

COV: Covariance

CV: Coefficient of variation

DL: Dong Lam, Thai Binh, Vietnam

EDL: Electrical double layer

F: Formation factor

FR: Friedersdorf, Germany

G: Gibbsite sheet

LN: Loebnitz, Germany

MDS: Multidimensional scaling

NX: Ngo Xa, Vietnam

PCA: Principal component analysis

SE: Standard error

SVD: Singular value decomposition

TL: Tra Linh, Vietnam

VD: Vu Doai, Vietnam

XRD: X-ray diffraction

YD, Yen Dinh, Vietnam.

Symbols

$\vec{B}$: Magnetic induction vector (Tm^{-2})

c: Cohesive force (kPa)

C_c: Coefficient of curvature

$\vec{D}$: Dielectric displacement vector (Cm^{-2})

δ_{rs}: Distance between two points r and s

e: Void ratio

$\vec{E}$: Electric field vector (Vm^{-1})

ϵ: dielectic permittivity (Fm^{-1})

ϵ_r: Relative dielectric permittivity

ϕ: Porosity (%)

φ: Internal friction angle (^{0})

G_s: Specific gravity

$\vec{H}$: Magnetic field vector (Am^{-1})

$\vec{J}$: Conduction current density vector (Am^{-2})

k: Hydraulic conductivity (m/s)

$\mathfrak{L}$: r-dimensional subspace

LL: Liquid limit (%)

μ: Magnetic permeability (Hm^{-1})

PI: Plasticity index (%)

PL: Plastic limit (%)

ρ: Electrical resistivity (Ωm)

σ: Electrical conductivity (Sm^{-1})

σ_f: Normal stress on the failure plain (kPa)

σ_i: Eigenvalues

$\mathbf{R}$: Variance-covariance matrix

$\mathbf{\Sigma}$: Variance-covariance matrix

SSA: Specific surface area (m^2/g)

τ_f: Shear strength on the failure plain (kPa)

$\mathbf{U}$: Unitary matrix

$\mathbf{V}$: Unitary matrix

w: Gravimetric water content (%)

w^*: Volumetric water content (%)

$\mathbf{x}$: vector of a data set

$\mathbf{X}_{n\times p}$: Data matrix of n rows and p columns

$\tilde{\mathbf{X}}$: Standardized data matrix

$\hat{\mathbf{X}}_r$: Approximation of $\mathbf{X}$ in a lower r-dimensional subspace.

$\mathbf{Y}$: Score matrix.

Contents

List of Figures

List of Tables

Chapter 1

Motivation

Vietnam belongs to the countries possessing the longest dike systems in the world. The country has a very extensive dike system, mainly located in Red River Delta in the northern part, including 8,000 km of river dikes and 2,000 km of sea dikes. Being a country of monsoonal climate with high rainfall, high water events and frequent typhoons make the country vulnerable to severe flooding. Since the first dike was set up in 11^{th} century, the dike system has been gradually and intermittently built (Tinh, 2001). The existing river dikes were built long ago by hand using local materials and without much knowledge on geotechnical principles. Because earth fills and foundations were neither selected nor treated carefully, there are frequent sand boiling, piping seepage and slides along nearly every section of dikes. Thus, during large floods of long duration, there occur dike breaches and embankment failures, especially where the dikes have poor foundations and water ponds on both sides because of unauthorized earth excavations in the past. Water level fluctuation in the river channels during floods can cause bank erosion. Changes of soil behavior in the dike body as well as in its foundation are threatening the safety of the dikes. In recent years, geophysical methods in investigation of dikes and embankment dams have increasingly been applied. Weller et al. (1996) and Tuyen et al. (2000) have successfully applied vertical electric sounding (VES), electrical tomography

and well logging methods to detect permeable and fracture zones in the dike body and the foundation of dikes in Hanoi. Termite nests in dike body and defective sluices under the dike are dangerous threats of dike stability during flood season. These problems have been well studied by multi-electrode measurement in Thai Binh dikes (Weller et al., 2006). The combination of geophysical methods i.e. electrical tomography, seismic refraction and well logging has proved to be a useful tool in characterizing dike structures and identifying weak zones in the dike foundations in Nam Dinh and Thai Binh provices (Canh et al., 2005). For purpose of high-efficiency exploration of the river embankment, Fauchard and Mériaux (2007), Takahashi and Yamamoto (2010) suggested geophysical methods such as Slingram (low-frequency near-field electromagnetic method), radio magnetotellurics (low-frequency, far-field electromagnetic method), ground penetrating radar (GPR), electrical imaging and seismic refraction. Some in-situ geotechnical methods described as easy-to-use and effective as penetrometric tests (CPT, DCPT), Lefranc permeability tests, shear tests with phicometer and mechanical drilling were proposed to use in addition to geophysical methods mentioned above. In investigation of dike system in Germany, Weller et al. (2008), Niederleithinger et al. (2008) proposed geophysical methods like geoelectric, electromagnetic, ground penetration radar, surface wave seismic methods and engineering geophysical soundings. Interpretation and comparisons among them were made with assistance of core drilling results. The geophysical methods of engineering geophysical sounding and geoelectric methods have proved to be the most effective. For safety evaluation of embankment dams, internal erosion and leakage are the major problems. Johansson and Dahlin (1996) observed variations of resistivity and temperature by monitoring potential seepage in embankment dams. Phenomena of internal erosion and seepage induced seasonal variation inside the embankment dam were studied in the hydro-power dam of Hällby in Sweden (Sjödahl et al., 2008).

In investigation of dikes and embankment dams, non-destructive geophysical methods are often preferred since geotechnical methods like drilling and other penetrating investigations are normally not allowed. Some previous studies have been attempted to find relations between geotechnical and geophysical parameters on different sites (Braga et al., 1999; Ayres and Theilen, 2001; Giao et al., 2003; Ngoc, 2005; Cosenza et al., 2006; Schwartz et al., 2008 and Sudha et al., 2008). The study of the geotechnical and geophysical properties of soils and their relations actually requires an interdisciplinary scientific approach that is associated with an involvement of mechanics (loading) as well as the response to the fluctuations in the local environmental conditions. Local environment here refers to ambient conditions that are reflected by such variables as temperature, pressure, groundwater table and composition, microbial population, etc. that may change the engineering behaviors of soil. In the foundation, the soils should be considered as a multiphase soil-water-air system with varying degrees of water and air due to variation of groundwater table and a variety of physical and chemical processes that occur within these phases.

In this study, various types of soils from the dike systems in Vietnam and Germany were collected. The petrophysical properties i.e. complex conductivity, complex dielectric permittivity, magnetic susceptibility, grain density were investigated in laboratory in Germany. The geotechnical properties of soil such as grain size distribution, water content, Atterberg limits, density and strength parameters were measured in laboratory in Vietnam. Diep et al. (2004) reported that the instability of dike in Vietnam comes from the dike foundation rather than the dike body due to a high clay content in the dike materials. Hydraulic conductivity is usually a vital parameter in assessing the earth work's stability. For the clayey soils, soil permeability is less significant and it takes time to perform permeability tests. However, some soil samples at the dike monitoring system in Thai Binh, Vietnam, were collected and permeability test were performed. In the field, the dike monitoring system including

a multi-electrode array across the dike, sensors of tensionmeter and frequency domain reflectometery (FDR) was installed at a dike section in Thai Binh province to monitor the variations of resistivity, water content, temperature, dielectric permittivity with varying water table in the Red river.

A comprehensive set of soil samples originating from dike body and dike foundation is investigated in this study. The main aims of the thesis are:

- Investigation of petrophysical and geotechnical properties of soils from various locations in Vietnam and Germany;
- Identification of relationships between petrophysical and geotechnical parameters;
- Analysis of petrophysical and geotechnical properties of soils using multivariate statistics.

Chapter 2 describes the nature of soil and soil water interaction. Clay mineralogy of soils, geotechnical and petrophysical properties of soils are reviewed. The methods to determine geotechnical and petrophysical parameters in laboratory are also described in this chapter.

Chapter 3 shows the origin of soil samples collected in dike systems in Vietnam and Germany. The results of the mineralogical, geotechnical and petrophysical investigations are presented and discussed.

Chapter 4 deals with the theory of multivariate statistics tools in brief. Two multivariate statistic tools, namely principal component analysis and cluster analysis are presented in the chapter. The application of multivariate statistic tools to investigate petrophysical and geotechnical properties of soils including results and discussions are presented in chapter 5. Conclusions and recommendations are given in Chapter 6.

Chapter 2

Geotechnical and petrophysical characterization of soils

2.1 Introduction

From civil engineering point of view, soils are aggregates of mineral particles, organic matter and the void spaces between the particles containing water and/or air. Soils are special engineering materials mostly characterized by inhomogeneity and anisotropy. The resulting geotechnical and petrophysical properties vary over a wide range. The reason is that soils are natural materials, formed by the weathering of rocks. The behavior of soils is a legacy of natural processes, from their origin to the actual state. The discrete particles that make up soils are not strongly bounded together, they are free to move relatively among themselves and, when a soil element deforms, the overall deformation is essentially the result of relative sliding between particles and rotation of particles (Mitchell and Soga, 2005). Therefore, it is obviously that soil behavior is highly non-linear and irreversible.

In general, soils can be grouped into categories coarse and fine-grained or granular and cohesive, respectively. Coarse-grained soils have a macro structure with larger

size soil particles such as cobble, gravel and sand. Fine-grained soils have a micro structure that includes silt and clay. The characteristics of granular soils such as the grain size distribution and the grain shape affect the engineering behavior of these soils. For fine-grained soils the grain size distribution has relatively little influence on the engineering behavior, but the interaction between water and fine grains significantly affects their behaviors.

The term of clay sometimes makes ambiguous sense. When referring to particle size term, it indicates all constituents of a soil smaller than some given sizes, usually 2 μm or 5 μm as defined under various soil classification systems. As a mineral term, it refers to specific minerals termed clay minerals. In geotechnical engineering, the term clay is also used to describe clay soil - a soil, which contains clay minerals as well as other mineral constituents, has plasticity and is cohesive. In this sense a clay soil may contain little more than 10% of clay-size particles. It is so described because the presence of clay size particles significantly affects the behavior of soil. Thus, in order to avoid confusion, it is useful to use the term clay size when referring to compositions in terms of particle size and clay mineral content or simply clay content when speaking of clay mineral compositions.

2.2 Soil mineralogy

Mineralogy is the primary factor controlling the size, shape, and physical and chemical properties of soil particles. Therefore, a knowledge of soil mineralogy is essential to a fundamental understanding of soil behavior. Clay minerals are very tiny crystalline substances evolved primarily from chemical weathering of certain rock-forming minerals. Almost all clay minerals are very small in micro-sized crystals. They are complex silicates containing also aluminum, magnesium, iron and potassium. Two

basic crystalline units form the clay minerals: (i) a silicon-oxygen tetrahedron, and (ii) an aluminum or magnesium octahedron (Grim, 1962).

A silicon-oxygen tetrahedron unit, shown in Figure 2.1a, consists of four oxygen

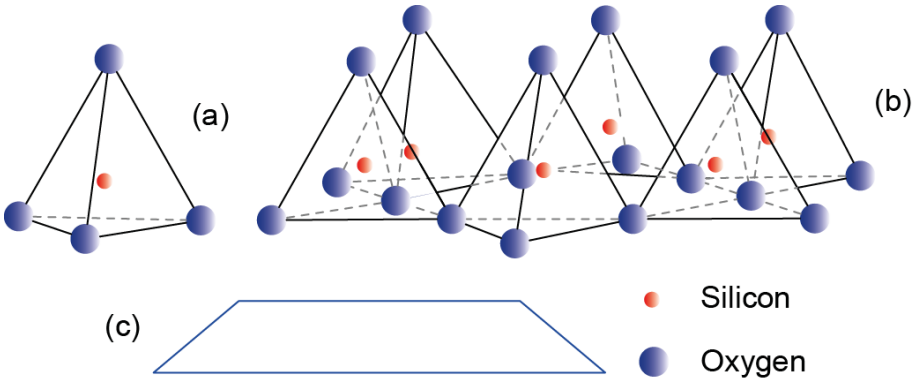


Figure 2.1: Sketch of basic structural unit in tetrahedral sheet: (a) Silica tetrahedron, (b) Silica tetrahedral sheet, and (c) Symbol of a silica sheet (modified after Das, 2008).

atoms surrounding a single silicon atom. The oxygens at the base of each tetrahedron are in one plane, and the unjoined oxygen corners all point in the same direction. The tetrahedral units combine to form a silica sheet as shown in Figure 2.1b. The three oxygens located at the base of each tetrahedron are shared by neighboring tetrahedra. Each silicon with a positive valence of four is linked to four oxygen atoms with a total negative valence of eight. The structure can repeat indefinitely and has the composition $(Si_4O_{10})^{4-}$. This leaves one negative valence charge of the top oxygen of each tetrahedron to be counter balanced. The electrical neutrality can be obtained by substitution of four oxygens by hydroxyls or by union with another sheet of different composition that is positively charged. The oxygen-to-oxygen distance in the silica tetrahedral sheet is 2.55 Å, the space available for the silicon ion in tetrahedral coordination is 0.55 Å, and the thickness of the sheet in clay mineral structures is 4.33 Å (Grim, 1962).

An octahedral sheet basically is a combination of octahedral units enclosing an aluminum, magnesium, or other cation. Figure 2.2a shows a single octahedron

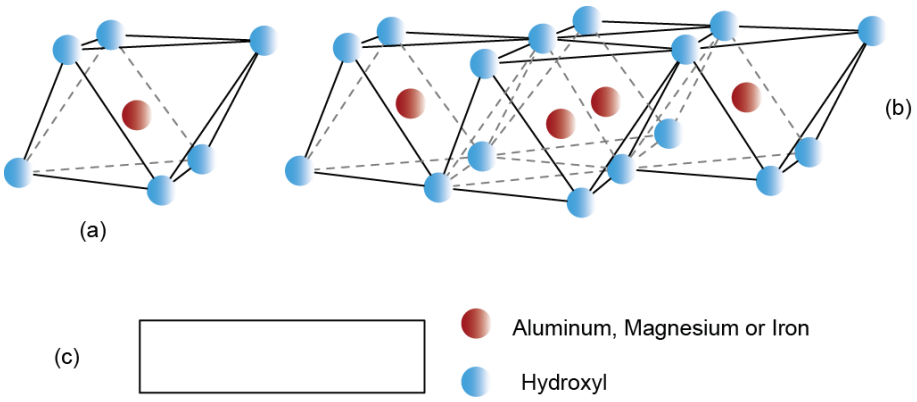


Figure 2.2: Sketch of basic structural unit in octahedral sheet: (a) Single octahedral unit, (b) Octahedral sheet, and (c) Symbol of a octahedral sheet (modified after Das, 2008).

consisting of six hydroxyls surrounding an aluminum (or a magnesium) cation, while Figure 2.2b illustrates how octahedra combine to form a sheet structure. In some case, other cations are present in place of Al^{3+} and Mg^{2+}, such as Fe^{2+}, Fe^{3+}, Mn^{2+}, Ti^{4+}, Ni^{2+}, Cr^{3+}, and Li^{+} (Lambe and Whitman, 1969). The substitution of different cations in the octahedral sheet is rather common and results in different kind of clay minerals. Since the cations substituted are at approximately the same physical size, such substitution is called isomorphous. The combination of the aluminum octahedral units forms a gibbsite sheet symbolized as G. If the main metallic cations in the octahedral units are magnesium, this sheet is called brucite sheet symbolized as B. The oxygen-to-oxygen distance is 2.6 Å, the OH-OH distance is 2.94 Å and the space available for cations is 0.61 Å. The thickness of the sheet is 5.05 Å (Grim, 1962).

The variations in the basic sheet structures make up thousands of clay minerals. All clay minerals consist of two basic sheets which are stacked together in certain unique way and with certain cations present in the tetrahedral and octahedral sheets. When the silica sheets are stacked over the octahedral sheets, as shown in Figure 2.3, the oxygen atoms replace the hydroxyls to satisfy their valence bonds. This sheet is about 7.2 Å thick. The repeating layers are held together by hydrogen bonding and

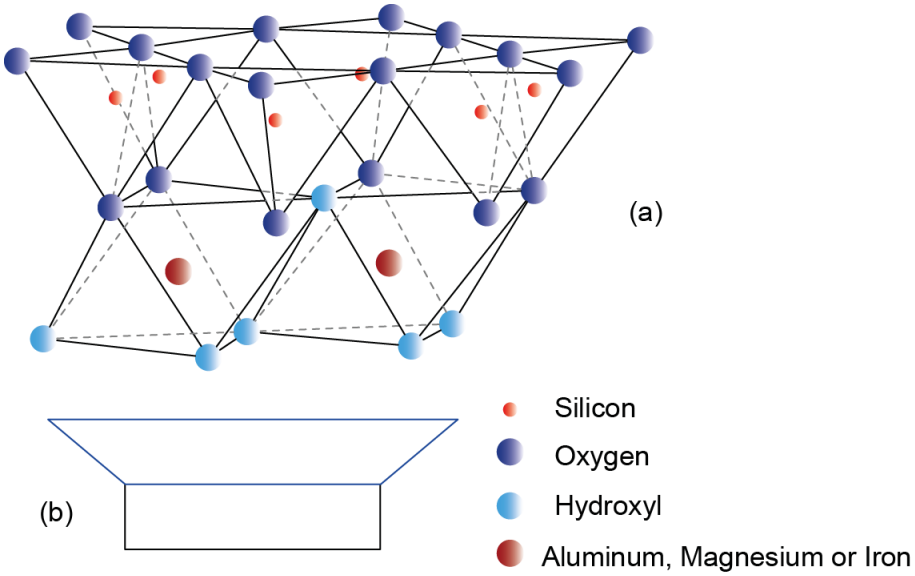


Figure 2.3: Structure of silica-octahedral sheet:(a) Elemental silica-octahedral sheet, and (b) Symbol of silica-octahedral sheet (modified after Das, 2008).

secondary valence forces.

For engineering purposes, it is usually sufficient to describe only a few of more common clay minerals which are found in clay soils.

Kaolinite consists of repeating layers of one silica sheet and one octahedral sheet. Because of the stacking of one layer of two basic sheets, kaolinite is called 1:1 or two-layer clay mineral as shown in Figure 2.4a. The two sheets held together in such a way to form a single layer with 7.2 Å in thickness and extends indefinitely in the other two directions. The successive layers of the basic layer are stacked together by hydrogen bonds between hydroxyls of octahedral sheet and the oxygens of tetrahedral sheet. Because the hydrogen bond is very strong, it is therefore unable to hydrate and allows the layers to stack up to make a rather large crystal. Yong and Warkentin (1975) reported that kaolinite is the largest, thickest clay mineral with a thickness of 0.05 - 2 μm. A typical kaolinite crystal can reach a thickness of 70 to 100 layers (Holtz and Kovacs, 1981).

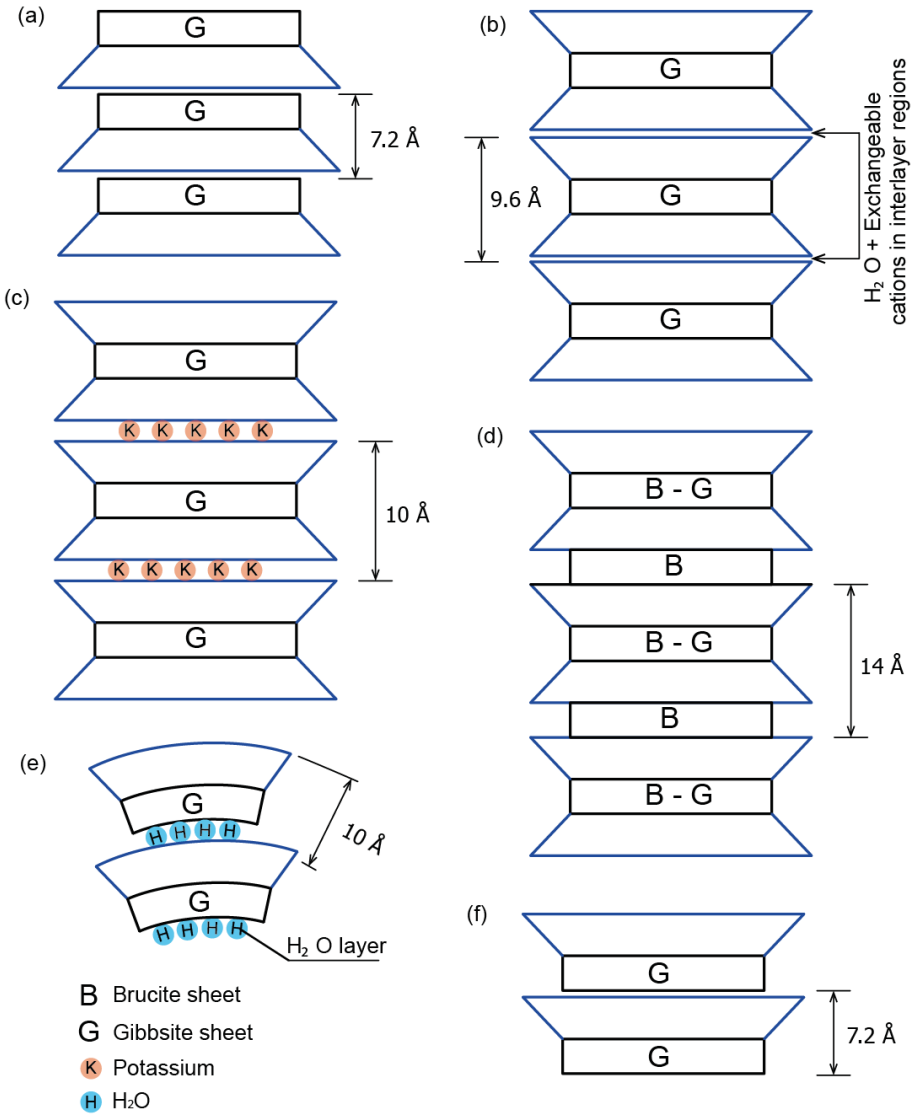


Figure 2.4: Diagrammatic sketch of the structures of some common clay minerals: (a) Kaolinite, (b) Montmorillonite, (c) Illite, (d) Chlorites, (e) Halloysite (10Å), and (f) Halloysite (7.2Å) (modified after Grim, 1962 and Mitchell and Soga, 2005).

Montmorillonite is an important mineral that consists of two tetrahedral sheets and one octahedral sheet, that is why montmorillonite is called a 2:1 or three-layer mineral as illustrated in Figure 2.4b. The octahedral sheet is sandwiched between two tetrahedral sheets where the oxygens at tips of tetrahedra combining with hydroxyls of octahedral sheet to form a single layer of approximately 9.6 Å thickness. Like kaolinite, the layers stack together and extend indefinitely in the other two

directions. The layers are stacked together by van der Waals bonds which are rather week compared to hydrogen bonds and there is a net negative charge deficiency in octahedral sheet. Water and exchangeable ions can occupy and separate the layers (Mitchell and Soga, 2005). Hence, a typical montmorillonite crystal is rather thin with the thickness of 30 Å as illustrated in Figure 2.5. But it has strong attraction of water and exchangeable ions. Soils containing montmorillonite are very susceptible to swelling when increasing their water content and the development of swelling pressure can easily damage light structures as earth works and embankment dikes. This is a major problem worldwide.

Illite is the most commonly found clay mineral in soils encountered in engineering practice. It also has the structure of 2:1 similar to montmorillonite, but the layers are bonded together by potassium cations. There are some isomorphous substitutions of aluminum for silicon in tetrahedral sheet and some illites may contain magnesium and iron in octahedral sheet as well as aluminium. Interlayer bonding by potassium is so strong that the basal spacing of illite remains fixed at 10 Å as schematically shown in Figure 2.4c. Illite usually occurs as a very small flaky particle mixed with other clay minerals. A typical flake thickness may be as small as 300 Å.

Chlorite, relatively common in clay soils, has a basic four-layer structure of 2:1:1 consisting of two tetrahedral sheets bonding by a gibbsite or brucite sheet in between and a brucite outside as conceptually illustrated in Figure 2.4d. The basal spacing is fixed at 14 Å. It is basically similar to illite except that an organized octahedral sheet replaces the area otherwise occupied by potassium ions. Chlorite minerals occur as microscopic grains of platy morphology and poorly defined crystal edges with a typical thickness of 300 Å. In soils, chlorites always occur in mixtures with other clay minerals.

Halloysite is a form of kaolinite. The basal spacing for the nonhydrated form is about 7.2 Å, as for kaolinite. Because of the connecting water layer, the basal spacing

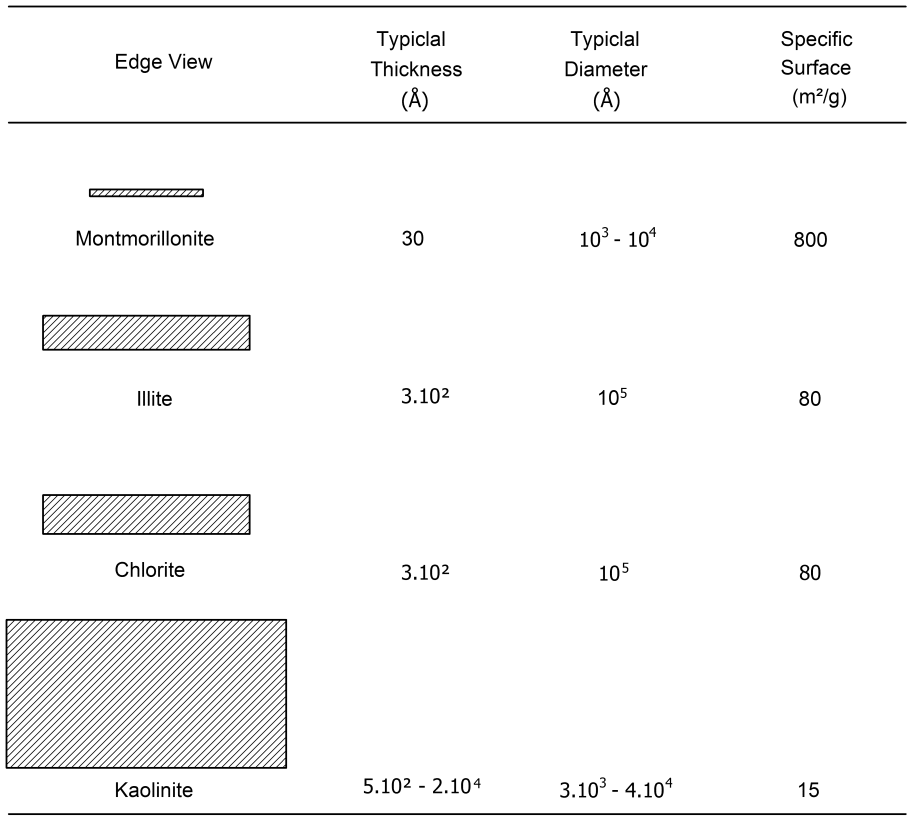


Figure 2.5: Average thicknesses, diameters and internal surfaces of common clay minerals (after Yong and Warkentin, 1975 and Holtz and Kovacs, 1981).

for hydrated halloysite is about 10.1 Å. The difference between these values, 2.9 Å, is the approximate thickness of a single layer of water molecules. The recommended terms for the two forms are halloysite (10.1 Å) and halloysite (7.2 Å) as shown in Figure 2.4e and 2.4f, respectively. An irreversible transformation from halloysite (10.1 Å) to halloysite (7.2 Å) by dehydration can occur at approximately 50^0 C. Halloysite shows plates with curled edges with a wall thickness of 200 Å and it is often found in soils formed from volcanic parent materials in wet environments.

The structures in detail and development of other clay minerals in environmental sedimentations are well documented in Paquet and Clauer (1997), Meunier (2005), Murray (2007) and Velde and Meunier (2008).

The most widely used method of identification of clay minerals is from an X-Ray Diffraction (XRD) pattern of a soil sample of the clay-sized fraction. An

experiment performed with X-rays establishes the three prevailing concepts of X-ray diffraction: (i) atomic particles within crystals are arranged in orderly, three-dimensional, repeating patterns; (ii) these regular arrangements have spacings of approximately the same dimensions as the wavelength of X-rays and therefore (iii) X-rays are wavelike in nature (Moore and Reynolds, 1997). The equation relating the lattice spacing between planes to the monochromatic X-ray wavelength, known as Bragg's Law, is stated as

$$2dsin\theta = n\lambda, \tag{2.1}$$

where d is the lattice spacing between planes of atoms, θ is the angle of incidence between glass slide and x-ray beam, n is integral number (1, 2, 3...) relating to wavelengths, and λ is X-ray wavelength.

The soil samples were prepared using a qualitative analysis of clay minerals. The organic matters in soil samples were removed by mixing with water and Hydrogen Peroxide (H_2O_2 3%) solution. To prevent flocculation, a diluted NH_3 solution was used as deflocculant. The soil samples were then centrifuged to separate out the fraction smaller 2 μm using a programmed centrifuge with three-minute runtime at 1,037 revolutions per minute (rpm) including acceleration and deceleration. Oriented slides were prepared using the glass slide method. The detailed description of sample preparation can be found in Moore and Reynolds (1997) and Pansu and Gautheyrou (2006).

The X-ray diffraction was performed using Philips PW-1800 X-ray diffractometer with the X-ray generator of 40 kV and 30 mA and Cu Kα radiation.

2.3 Soil water interaction and water clay electrolyte system

As mentioned above, fine grained soils, especially clay soils, are strongly influenced by the presence of water. The amount of water existing in the fine grained soils will significantly influence the engineering behavior of soil. The interaction between water and soil can basically occur only at the soil mineral surface. The mechanism of soil water interaction is complex and its behavior is not only dependent on soil types, but is also related to the current and past environmental conditions and stress history.

Clay particles in soils are always hydrated, i.e. surrounded by layers of water molecules called adsorbed water or hygroscopic water. Mitchell (1993) described that two water layers are extremely well arranged around negatively charged clay surfaces. The first is called tightly bound layer with a thickness of 20 Å and the outer layer is called loosely bound layer with a thickness of 20 - 60 Å. These water molecules should be considered as a part of the clay surface when the behavior of clay soils is considered.

Plasticity, compaction, inter-particle bonding and water movement in soils are all influenced by the water layers. The properties of clays change as the thickness of this hydration shell changes, and consequently the engineering properties of soils change as well. The water in soil which is neither bound around clay surface like adsorption water nor responds to gravity like free water, is usually called viscous water or capillary water. Figure 2.6 illustrates the distribution of adsorption water and capillary water in loose and compacted soil aggregates. The force holding water molecules to the clay surface arises both from the water and the clay. Water is a dipole molecule with a separation of centers of positive and negative charge. That

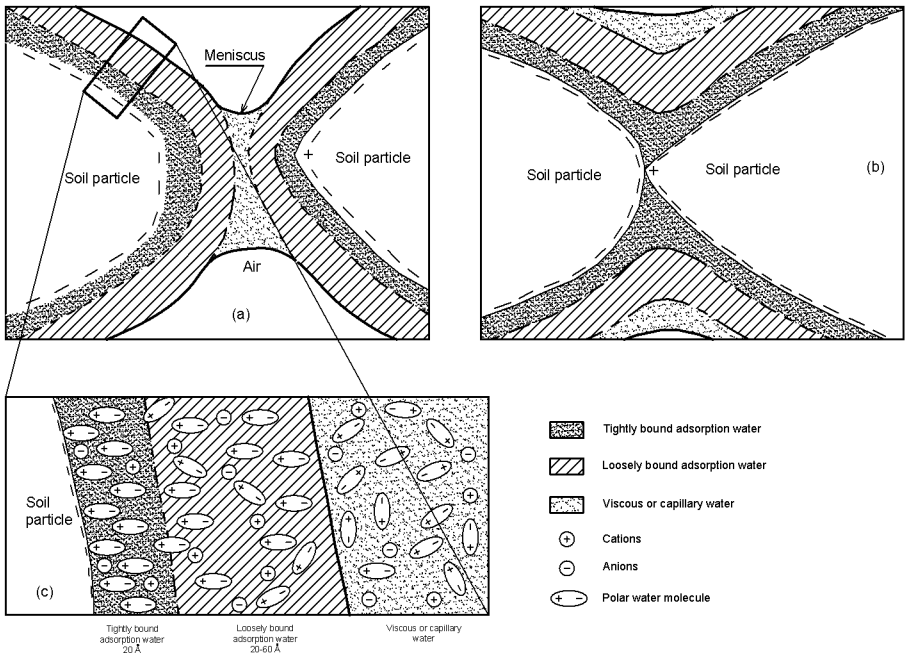


Figure 2.6: Sketch of bound water surrounding clay soils: (a) Loose structure, (b) Compacted structure, and (c) Adsorption and capillary water at clay surfaces (modified after Mitchell, 1993 and Saarenketo, 1998).

is why water is attracted by charges on the clay surface. The main force bonding water to the surface is due to the hydrogen bond. The first layer of water molecules is held by hydrogen bonding to the clay surface. The second water layer is held to the first, again by hydrogen bonding, but the force becomes weaker with distance as the orientating influence of the surface on the water molecules decreases. Each successive layer is held less strongly and the bonding quickly decreases to that of free water.

The substitution of one ion for another in the clay crystal lattice is often found as described in previous section. Imperfections at the surface consequently occur, especially at the edges. That leads to negative electric charges on clay particles. Cations from the pore water are attracted to the particles and anions are repelled to guarantee electroneutrality. These are the exchangeable cations and their number is the cation exchange capacity (CEC) or the amount of negative charge per unit weight or per unit surface area of the clay. Another source of electric charge on clay

particle is the unbalanced valence charges at the edges of the particles as known as broken-bond charges (Mitchell, 1993).

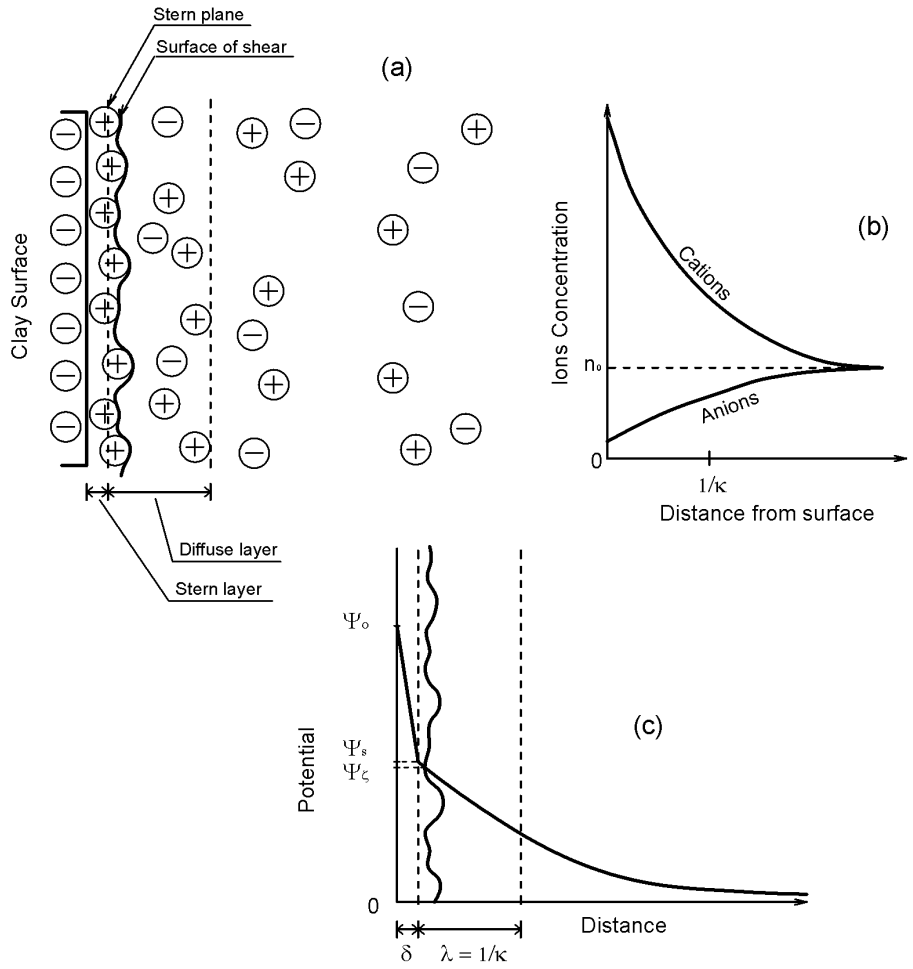


Figure 2.7: The structure of Electrical Double Layer (EDL) adjacent to clay surface - Stern-Gouy model: (a) Distribution of ions and models of Stern and diffuse layers; (b) Distribution of cations and anions from clay surface; and (c) Potential distribution according to Stern's theory (modified after Olphen, 1977; Shaw, 1992 and Mitchell and Soga, 2005).

The ion adsorption at clay particle surface can be explained by the Electrical Double Layer (EDL) theory. The electrical force between negatively charged surface and positively charged ions attracts the cations to the surface, but their thermal energy makes them diffuse away from this space with a high ion concentration. The balance of Coulomb electrical attraction and thermal diffusion results in a diffuse

layer of cations, with the highest concentration at the surface and gradually decreasing with distance from the surface. The theory of electrical double layer deals with the distribution of these ions and the magnitude of the electrical potentials which result in the vicinity of the charged surface (Olphen, 1977). The interaction of diffuse ion layers of adjacent particles gives an explanation for the properties of swelling, plasticity, and water retention of clays. The electric double layer can be regarded as coexistence of two regions: (i) an inner region which may include adsorbed ions, and (ii) a diffuse region in which ions are distributed according to the influence of electrical forces and random thermal motion as shown in Figure 2.7a. From electrostatic and diffusion theory, the Poission-Boltzmann equation presents the distribution of positive and negative ions as a function of the distance from the surface as illustrated in Figure 2.7b. Both cations and anions reach equal concentration at a large distance from the surface. The area within cations and anions curves represents the total net diffuse layer charge equivalent to the surface charge. According to Stern model, the surface or wall potential changes from Ψ_0 at the surface to Ψ_s (the Stern potential) in the Stern layer and then decays exponentially to zero in the diffuse double layer as shown in Figure 2.7c. For practical purpose, the zeta potential Ψ_ζ, the potential between the surface of shear and bulk solution, is assumed to be identical to Stern potential Ψ_s. The thickness of Stern layer δ and mobile thickness of diffuse double layer $1/\kappa$, also known as Debye screening distance, are theoretically well documented in Shaw (1992) and Revil and Glover (1997).

2.4 Geotechnical properties

The engineering properties of a soil depend on the composite effects of several interacting factors. These factors may be divided into two groups: compositional

factors and environmental factors. Compositional factors determine the potential range of values for any property. They include types of minerals, amount of each mineral, types of adsorbed cations, shapes and size distribution of particles, pore water composition and type and amount of other constituents, such as organic matter, silica, alumina, and iron oxide. The influences of compositional factors on engineering properties can be studied using disturbed samples. Environmental factors determine the actual value of any property. They include water content, density, confining pressure, temperature, fabric and availability of water. Undisturbed samples, or in situ measurements, are required for the study of the effects of environmental factors on properties.

Study of soil properties requires inter-disciplinary approaches. The basic definition in geotechnical engineering may be found in many soil mechanics text books (Terzaghi et al., 1996; Aysen, 2002; Das, 2007). Some basic definitions in brief and involved soil properties are described here in an attempt to present a systematic approach and to avoid any misunderstanding terms may rise between disciplinary fields.

2.4.1 Soil index properties

In general, soil may be considered as a three-phase system consisting of solid grains and inter-particle void space filled with liquid or gas or both. The solids are small grains of different minerals, while the voids can be filled with water, air or filled partly with both water and air. Figure 2.8a depicts the soil skeleton consisting of three phases and the phase diagram of idealized soil is shown in Figure 2.8b. The total volume V of the soil mass contains the volume of soil solids V_s and the volume of voids V_v. The volume of voids is made of the volume of water V_w and the volume of air V_a. On the other side, the corresponding masses of the phases are indicated, those are total mass M_t, mass of water M_w and mass of solids M_s. It should be noted that for the practical purpose the mass of air is assumed to equal zero.

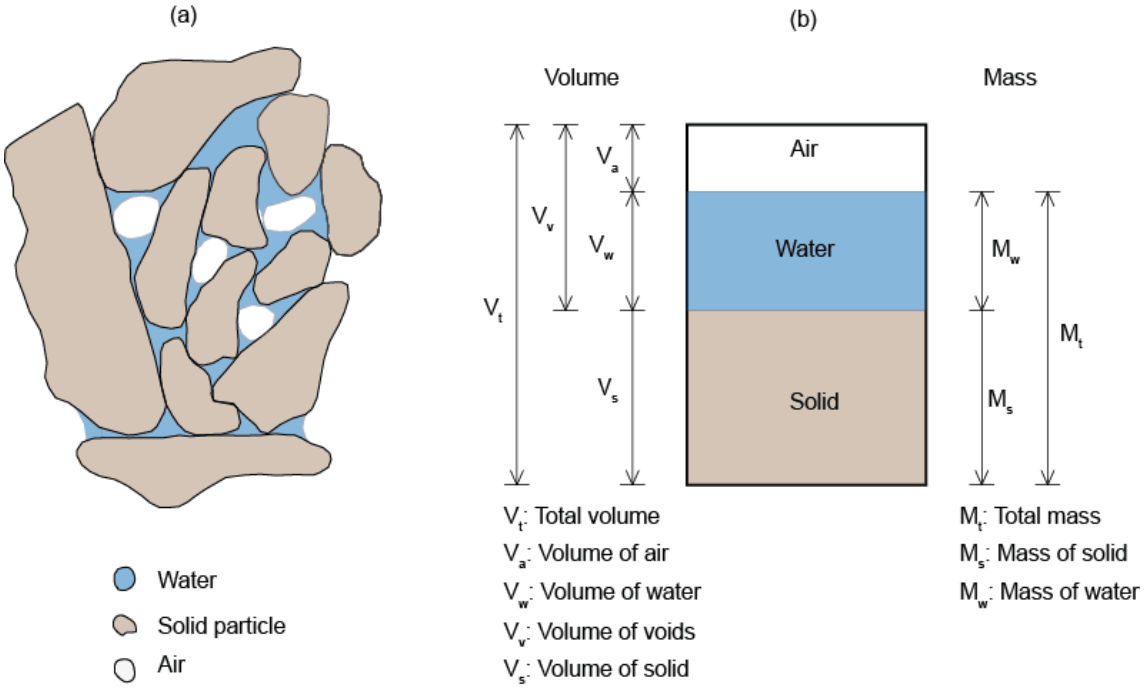


Figure 2.8: Components of air, water and solid in the soil mass: (a) Soil skeleton consisting of solid particles, voids and water (modified after Holtz and Kovacs, 1981) and (b) Phase diagram of idealized soil volume and soil mass.

Volumetric ratios

There are three volumetric ratios that are very useful in geotechnical engineering and these can be determined directly from the phase diagram.

The *void ratio (e)* is defined as

$$e = \frac{V_v}{V_s}, \tag{2.2}$$

where V_v is the volume of voids, and V_s is volume of the solids.

The void ratio e is normally expressed as a fraction.

The *porosity (ϕ)* is defined as

$$\phi = \frac{V_v}{V_t}, \tag{2.3}$$

where V_t is the total volume of the soil sample.

The porosity ϕ can be expressed as a fraction or a percentage.

The void ratio and the porosity are inter-related as

$$e = \frac{\phi}{1-\phi}, \tag{2.4}$$

$$\phi = \frac{e}{1+e}. \tag{2.5}$$

The *degree of water saturation* S_w is defined as

$$S_w = \frac{V_w}{V_v}, \tag{2.6}$$

where V_w is the volume of water.

It can be expressed as a fraction or a percentage. It indicates what percentage of the total volume of voids contains water.

Density

Another very useful parameter in geotechnical engineering is density which is generally expressed as mass per unit volume. There are several commonly used densities. These may be defined as the total or wet density, ρ_t; the dry density, ρ_d; and the density of solid particles or grain density, ρ_s as

$$\rho_t = \frac{M_t}{V_t}, \tag{2.7}$$

$$\rho_d = \frac{M_s}{V_t}, \tag{2.8}$$

$$\rho_s = \frac{M_s}{V_s}. \tag{2.9}$$

In the fields of soil science and geophysics, dry density ρ_d is also termed bulk density. The mass of oven dried sample is normalized to the total volume of the original

sample. Bulk density, grain density, and porosity are related by the following equation

$$\rho_d = \rho_s(1 - \phi) \tag{2.10}$$

that can be used to determine the porosity of a soil sample. Equation 2.10 reflects an inverse relation between bulk density and porosity: an increase porosity of a soil lowers its bulk density.

In the case of fully saturated soil, S_w=100%, the saturated density can be determined

$$\rho_{sat} = \frac{M_t}{V_t} = \frac{M_s + M_w}{V_t}. \tag{2.11}$$

The grain density in this study was determined by Ultrapycnometer equipment using helium gas which can penetrate the finest pores of soil mass. The volume and true density of solid particles were measured by employing the gas expansion method. Approximately 50 grams of soil sample were prepared and dried in the oven about 24 h. After calibration with the standard known-volume-density steel ball, volumes and grain densities are automatically calculated until reaching the desirable deviation of less than 0.02%.

Water content

The water in soil significantly determines the behavior of the soils. It is very helpful to know how much water is present in the voids relative to the mass of solids. The gravimetric water content w of a soil mass is defined as the ratio of the mass of water M_w in the voids to the mass of solids, M_s,

$$w = \frac{M_w}{M_s}. \tag{2.12}$$

In laboratory standard water content is usually determined in accordance with the standard test method for laboratory determination of water (moisture) content of soil and rock by mass (ASTM D2216-98, 1998). A test specimen was dried in an oven at a temperature of $110^0 \pm 5^0$C in 16 hours. The loss of mass due to drying is considered to be water. The water content was calculated using the mass of water and the mass of the dry specimen.

In geoscience, soil science or in the case of some theoretical considerations of water retention and flow in soils or calculations for irrigation and drainage, water content on a volume basis is widely used. The volumetric water content w^* represents the fraction of the total volume of soil that is occupied by the water contained in the soil and can be defined as volume of water per total volume of soil,

$$w^* = \frac{V_w}{V_t}. \tag{2.13}$$

The volumetric water content can be expressed in terms of the mass-basis water content as

$$w^* = w\frac{\rho_d}{\rho_w}, \tag{2.14}$$

where ρ_d and ρ_w are dry density and water density, respectively.

The volumetric water content is also expressed in terms of the porosity and the water saturation as

$$w^* = \phi \; S_w. \tag{2.15}$$

Specific gravity

Specific gravity of solids G_s is defined as the ratio of the weight of the solids to the

weight of an equivalent volume of water,

$$G_s = \frac{\rho_s}{\rho_w} = \frac{M_s}{V_s \rho_w}, \tag{2.16}$$

where ρ_w is the density of water at 4^0C equal to 1.00 g/cm^3. Since the variation in density is relative small over the range of temperature encountered in ordinary engineering practice, the density of water ρ_w at other temperatures may be taken the same as that at 4^0C.

Grain size and its distribution

Soil types, in general, are termed as sand, silt and clay by their grain size. However, the grain size range for these terms are slight different among disciplines: geoscience, soil science and practical geotechnical engineering. There are also slight differences in soil classification of different countries. Figure 2.9 shows the particle size ranges of some main soil classification systems such as those of American society of testing and materials (ASTM), American association for state highway and transportation officials (AASHTO), German institute for standardization (DIN), British standards institution (BSI), State standards of the Russian federation (GOST) as well as of Vietnamese directorate for standards and quality (TCVN). The system which is quite popular among geotechnical engineers is the ASTM's classification system also known as unified soil classification system.

The particle size distribution of a soil is normally presented as a curve on a semilogarithmic plot, the ordinates being the percentage by mass of particles smaller than the size given by the abscissa. The flatter the distribution curve the larger the range of particle sizes in the soil; the steeper the curve the smaller the size range.

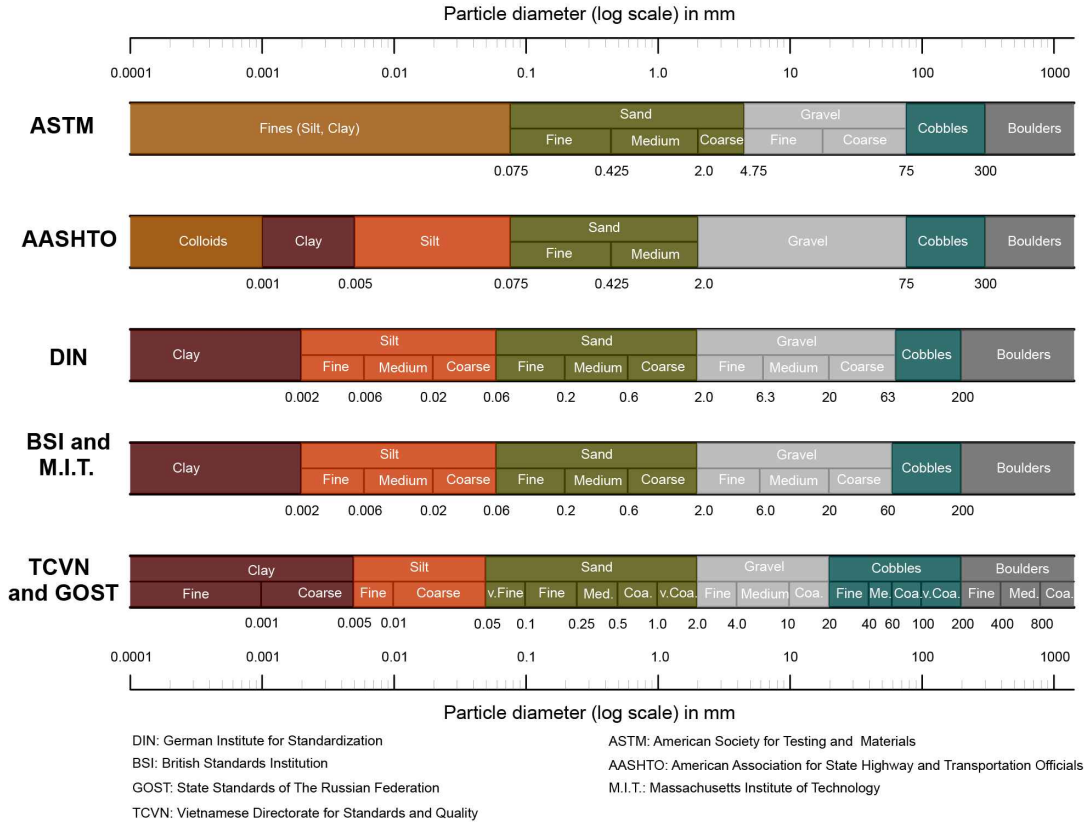


Figure 2.9: Grain size distributions associated with different engineering soil classification systems: ASTM D2487 (2000), AASHTO M145-91 (1991), DIN 18196 (2006), BSI BS EN ISO14688-2 (2004), GOST 25100-82 (1982), and TCVN 4198:1995 (1995).

A coarse soil is described as well graded if there is no excess of particles in any size range and if no intermediate sizes are lacking. In general, a well-graded soil is represented by a smooth, concave distribution curve. A coarse soil is described as poorly graded (i) if a high proportion of the particles have sizes within narrow limits (a uniform soil) or (ii) if particles of both large and small sizes are present but with a relatively low proportion of particles of intermediate size (a gap-graded or step-graded soil).

Particle size is represented on a logarithmic scale so that two soils having the same degree of uniformity are represented by curves of the same shape regardless of their positions on the particle size distribution plot. The particle size corresponding to

any specified value on the "percentage smaller" scale can be read from the particle size distribution curve as shown in Figure 2.10.

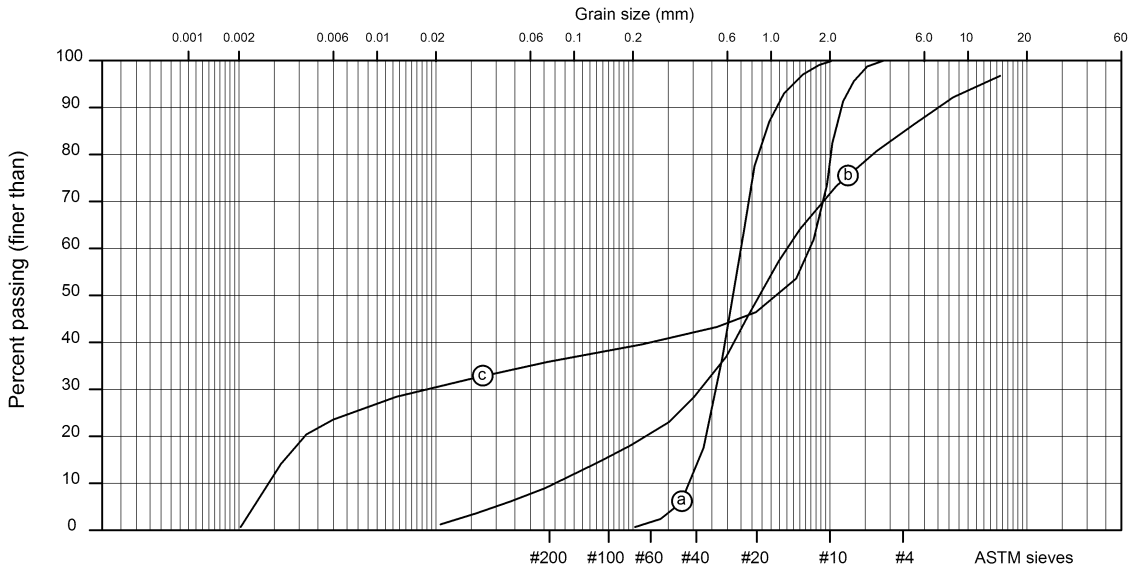


Figure 2.10: Particle size distribution curves: (a) Poorly or uniformly graded; (b) Well graded; and (c) Gap graded (after Holtz and Kovacs, 1981).

The size such that 10% of the particles are smaller than that size is denoted by D_{10}. Other sizes such as D_{30} and D_{60} can be defined in a similar way. The size D_{10} is defined as the effective size. The general slope and shape of the distribution curve can be described by means of the coefficient of uniformity (C_u) and the coefficient of curvature (C_c), defined as follows:

$$C_u = \frac{D_{60}}{D_{10}}, \tag{2.17}$$

$$C_c = \frac{D_{30}^2}{D_{10}.D_{60}}. \tag{2.18}$$

The higher the value of the coefficient of uniformity the larger the range of particle sizes in the soil. A well-graded soil has a coefficient of curvature between 1 and 3. In this study the determination of grain size distribution was performed in general

accordance with the standard test method for particle-size analysis of soils (ASTM D422-63, 1998). For coarse-grained soils, a sieve analysis is performed in which a sample of dry soil was shaken mechanically through a series of woven-wire square-mesh sieves with successively smaller openings. For fine grained soil, passes No. 200 sieve, a hydrometer analysis was conducted in a 1000 ml sedimentation cylinder with approximately 50 grams of soil. Hydrometer readings were taken at various elapsed times. Between readings, the hydrometer was removed from the soil suspension, rinsed and placed in the reference solution. The temperature and control fluid reading at each time interval was recorded.

The coefficient of uniformity C_u and coefficient of curvature C_c were, respectively, calculated as equations 2.17 and 2.18 using the values interpolated from the cumulative particle-size distribution curve.

Atterberg limits

Plasticity is an important characteristic in the case of fine soils. The term plasticity describing the ability of a soil to undergo unrecoverable deformation without cracking or crumbling. In general, depending on its water content, a soil may exist in one of the liquid, plastic, semi-solid and solid states as shown in Figure 2.11. The water content at which the transitions between states occur differ from soil to soil and are termed Atterberg limits.

The transition state from the liquid state to a plastic state is called the *Liquid Limit, LL*. At this stage, all soils possess a certain small shear strength. This arbitrarily chosen shear strength is probably the smallest value that is feasible to be measured in a standardized procedure. The transition from the plastic state to the semisolid state is termed the *Plastic Limit, PL*. The *Shrinkage Limit, SL*, is water content at which no more volume change occurs upon drying as illustrated in Figure 2.11.

The Atterberg limits determination was performed in general accordance with

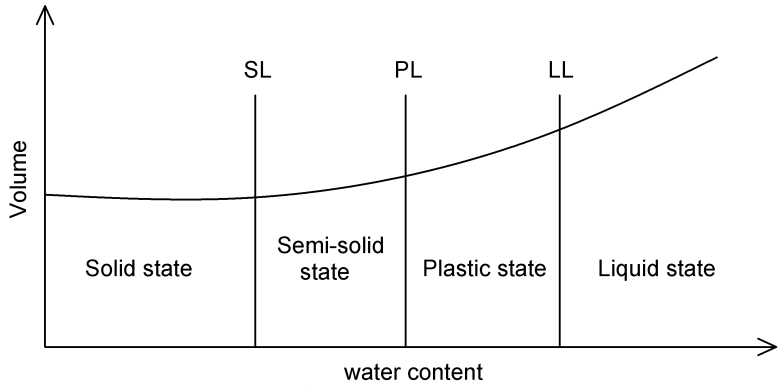


Figure 2.11: Atterberg limits relative to volume change and water content

the standard test method for liquid limit, plastic limit, and plasticity index of soils (ASTM D4318-00, 2000). Approximately 150 grams of soil are needed to complete the Atterberg limits test. Distilled water was added to the soil samples to bring the water content to a point where the blow count equaled 15 or less. The soil sample was covered and placed in the humid room to temper overnight. Approximately 20 grams of soil sample were set aside for the plastic limit determination and the rest was used for determining the liquid limit. Four separate water content determinations between 15 and 35 blows using the Casagrande cup were required to accurately determine the Liquid Limit of the particular soil sample. Once these data were plotted, the liquid limit was determined by locating the water content at 25 blows. For plastic limit determination, 1/3 of the 20 grams was taken and rolled into a 3.17 mm strand on the glass plate. This step was repeated until the soil crumbled when the soil reached 3.17 mm diameter. A water content determination was then performed. Two more strands were rolled and the water content determinations were conducted. The average of the three water contents was taken as the plastic limit.

The range of water content between the liquid and plastic limits, which is an important measure of plastic behavior, is called the *Plasticity Index, PI*,

$$PI = LL - PL. \tag{2.19}$$

The relative consistency of a cohesive soil can be defined by a ratio called the *Liquidity Index, LI*. It is defined as

$$LI = \frac{w - PL}{PI}, \tag{2.20}$$

where w is natural water content.

Significant values of liquidity index indicating the consistency of the soil are: $LI < 0$ indicating semi-plastic solid or solid state, $0 < LI < 1$ presenting plastic state and $LI > 1$, the soil in liquid state.

Plasticity chart

By using Atterberg limits, fine-grained soils can be classified according to plasticity chart as shown in Figure 2.12. The chart is divided into five regions by the vertical line $LL = 50\%$, two horizontal lines $PI = 4\%$ and $PI = 7\%$ and two inclined lines, U-line and A-line (ASTM D2487, 2000). The plasticity chart can also be used to

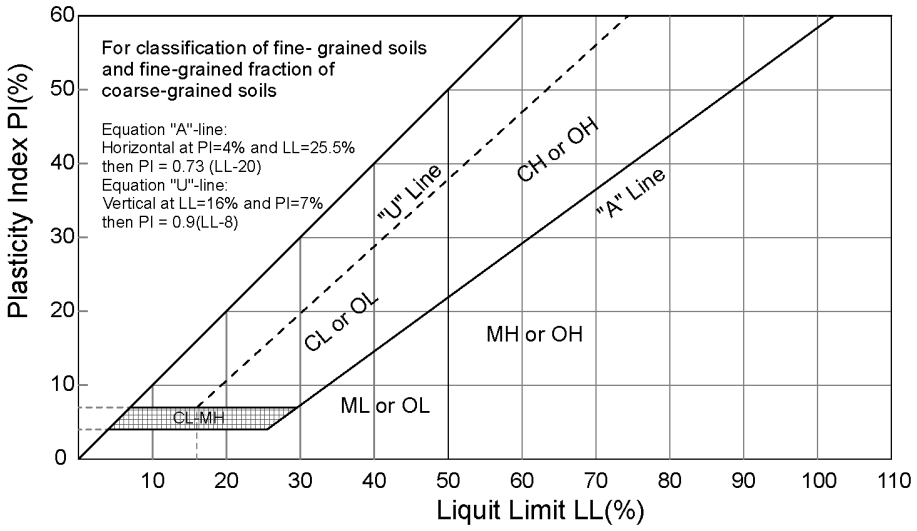


Figure 2.12: Plasticity chart: CL, lean clay; OL, organic silts, organic silty clay; CH, inorganic clay, fat clay; OH, organic clay; ML, inorganic silt, silty or clayed fine sand; MH, inorganic silt, elastic silt (ASTM D2487, 2000).

have a preliminary qualitative identification of the predominant clay minerals, by comparing the location of the tested sample with those of known minerals (Holtz and Kovacs, 1981). Figure 2.13 shows some main clay minerals overlaid on plasticity

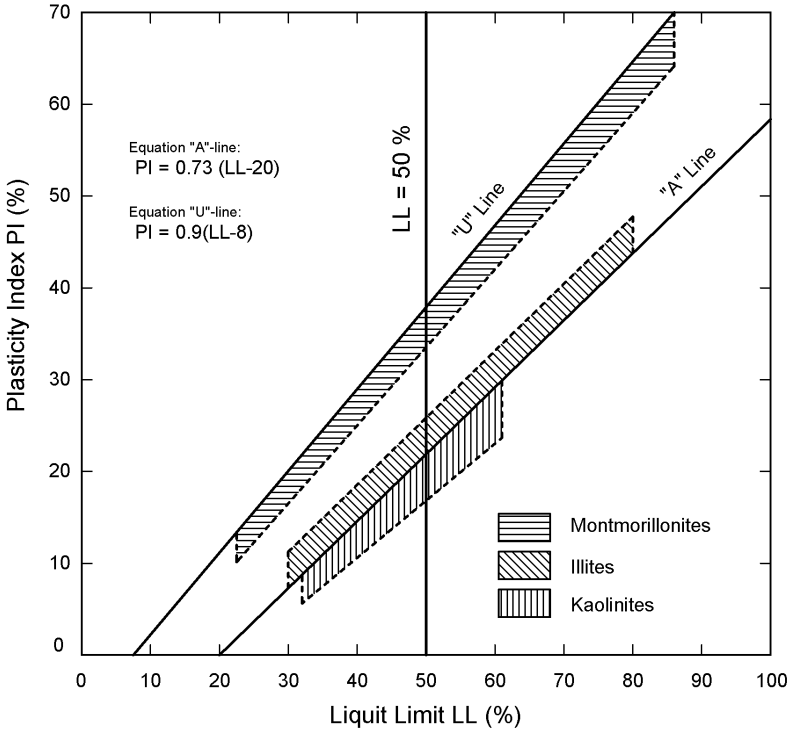


Figure 2.13: Clay minerals over plasticity chart (modified after Holtz and Kovacs, 1981 and Mitchell, 1993).

chart. If the tested sample plots near the U-line, the clay fraction is predominantly montmorillonite. Illites are located just above the A-line. Inactive kaolinite plot just below the A-line and their expected behavior is that of silts of medium to high compressibility. If the location plots to the left of the U-line, Atterberg limits values should be further checked.

2.4.2 Shear strength of soil

The shear strength of a soil mass is the internal resistance per unit area that the soil mass resists failure and sliding within any plane. Considering of the slope stability of embankment, bearing capacity of foundation and lateral pressure on earth-retaining structures,the shearing resistance of soil needs to be fully understood.

For most soil mechanics problems, it is sufficient to approximate the shear stress on the failure plane as a linear function of a normal stress. The linear function, well

known as *Mohr-Coulomb failure criterion*, can be defined as

$$\tau_f = \sigma_f tan\varphi + c, \tag{2.21}$$

where c is cohesion, φ is angle of internal friction, σ_f is normal stress on failure plane, and τ_f is shear strength on the failure plane.

The detailed theory of the criterion and inclination of failure plane caused by shear strength is well documented in many text books on soil mechanics such as Lambe and Whitman (1969), Holtz and Kovacs (1981), Terzaghi et al. (1996) and Das (2008).

In general, cohesion of soil depends on size, shape and clay mineralogy of the soil particles. Cohesion can be seen as the bonds between soil particles. For clay soil, cohesion is controlled by electrostatic forces which are attractive forces between particles and lubrication by water. Apparent cohesion is produced by capillary forces and interlocking friction of particle surfaces. The angle of internal friction can be seen as the angle at which the particle begins slide down the surface. Soil friction angle is controlled by porosity, particle size distribution and particle shape. Friction angle decrease with plasticity and water content.

There are several methods available to determine the shear strength parameters of various soil specimens. The direct shear test and triaxial test are widely used to determine strength parameters in laboratory. In this study the shearing properties of soil were measured by the direct shear test. The undisturbed soil sample is confined in the shear box of 20 mm in height and 60 mm in diameter. A vertical force is applied to the specimen through a loading plate and shear stress is gradually applied on the horizontal plane of the specimen. For each soil sample, three specimens with three different applied vertical forces were tested in undrained condition. The shear strength parameters were then determined from the best fitting line of shear stress versus normal stress.

2.4.3 Hydraulic conductivity

A soil is a permeable medium if liquid is free to move through the interconnected pore spaces between soil particles in accordance with Darcy's empirical law (Das, 2007)

$$v = \frac{q}{A} = ki, \tag{2.22}$$

where v is the discharge velocity, q is the volume of flowing water per unit time, A is the cross sectional area of the soil sample, k is the hydraulic conductivity and i is the hydraulic gradient. Hydraulic conductivity depends on the size of the interconnected pore spaces, particle shape and soil texture.

In the laboratory, hydraulic conductivity for soils can be determined by either constant head permeability test or falling head permeability test depending respectively on coarse-grained or fine-grained soil. The undisturbed soil specimen are confined in the sample holder ring of 20 mm in height and 60 mm in diameter. The water drains through the soil specimen into a water tank of constant level. Hydraulic conductivity is calculated from headwater lost after specific time of observation and geometry of the test equipment according to equation 2.22.

2.5 Soil classifications for engineering purposes

Different soils with similar properties may be classified into groups and sub-groups according to their engineering behavior. A classification system provides a common language to concisely express the general characteristics of soils, which vary infinitely without detailed descriptions. It provides a systematic method of categorizing soils according to their probable engineering behavior and allows engineers access to the accumulated experience of other engineers. A classification system does not eliminate the need for a detailed soil investigation or for testing for engineering properties.

However, the engineering properties have been found to be in good accordance with the index and classification properties of a given soil deposit. Thus, by knowing the soil classification, the engineer already has a fairly good general idea of the way the soil will behave in the engineering situation, during construction, and under loads. Since the properties of fine-grained soils can be correlated in a general way with the plasticity of the materials, classification systems for such soils are preferably based on the Atterberg limits rather than on grain size. Classification of mixed-grained soils containing both coarse and fine fractions should be based not only on the grain-size characteristics of the coarse fractions but also on the plasticity of the fine and very fine fractions.

There are several popular soil classification systems available and every system takes into consideration the particle-size distribution as such mentioned in Figure 2.9 and Atterberg limits. In spite of their insufficiencies, soil classifications based on grain-size characteristics are widely used, especially for preliminary or general descriptions. It is customary, in connection with such classifications, to assign the names of soils, such as "silt" or "clay" to different grain-size fractions. However, any system of classification based on grain size alone is likely to be misleading, because the physical properties of the finest soil fractions depend on many factors other than grain size. For example, according to anyone of the commonly used classifications, a soil consisting of quartz grains of colloidal size should be called a clay, whereas in reality it does not possess even a remote resemblance to clay. Hence, if the words "silt" or "clay" are used to express grain size, they should be combined with the word "size" as in the expression "clay-size particle" as mentioned in preceding section. The term fine is often used to describe the fraction of a soil that passes the No. 200 sieve (0.075 mm) and the term the clay-size fraction (CF) the fraction with sizes smaller than 0.002 mm or 0.005 mm.

The unsatisfactory nature of systems of soil classification based on grain size

Table 2.1: Unified soil classification system (USCS) according to ASTM D2487 (2000).

Major Divisions			Group of symbols	Typical Discriptions
COARSE-GRAINED SOILS More than 50% retained on the 0.075 mm (No. 200) sieve	**GRAVELS** 50% or more of coarsed fraction retained on the 4.75 mm (No. 4)	CLEAN GRAVELS less than 5% fines	GW	Well-graded gravels and gravel-sand mixtures, little or no fines
			GP	Poorly graded gravels and gravel-sand mixtures, little or no fines
		GRAVELS WITH FINES greater than 12 % fines	GM	Silty gravels, gravel-sand-silt mixtures
			GC	Clayey gravels, gravel-sand-clay mixtures
	SANDS 50% or more of coarse fraction passes the 4.75 mm (No. 4)	CLEAN SANDS less than 5% fines	SW	Well-graded sands and gravelly sands, little or no fines
			SP	Poorly graded sands and gravelly sands, little or no fines
		SANDS WITH FINES greater than 12% fines	SM	Silty sands, sand-silt mixtures
			SC	Clayey sands, sand-clay mixtures
FINE-GRAINED SOILS More than 50% passes the 0.075 mm (No. 200) sieve	**SILTS AND CLAYS** Liquid Limit 50% or less	INORGANIC	ML	Inorganic silts, very fine sands, rock flour, silty or clayey fine sands
			CL	Inorganic clays of low to medium plasticity, gravelly/sandy/silty/lean clays
		ORGANIC	OL	Organic silts and organic silty clays of low plasticity
	SILTS AND CLAYS Liquid Limit greater than 50%	INORGANIC	MH	Inorganic silts, micaceous or diatomaceous fine sands or silts, elastic silts
			CH	Inorganic clays or high plasticity, fat clays
		ORGANIC	OH	Organic clays of medium to high plasticity
Highly organic soils			PT	Peat, muck, and other highly organic soils

Prefix: G = Gravel, S = Sand, M = Silt, C = Clay, O = Organic
Suffix: W = Well Graded; P = Poorly Graded; M = Silty, L = Clay, LL < 50%; H = Clay, LL > 50%

alone led to a critical review of the problem (Casagrande, 1948) and the proposal of the Unified Soil Classification System (ASTM D2487, 2000). According to this system, which is presented in Table 2.1, all soils are divided into three major groups: coarse-grained, fine-grained, and highly organic (peaty). The boundary between coarse-grained and fine-grained soils is taken to be the 200-mesh sieve (0.075 mm). In the field the distinction is based on the decision whether the individual particles can be seen with the unaided eye. If more than 50% of the soil by weight is judged to consist of grains that can be distinguished separately, the soil is considered to be coarse-grained. The coarse-grained soils are divided into gravelly (G) or sandy (S) soils in accordance with the decision whether more or less than 50% of the visible grains are larger than the No. 4 sieve (4.75 mm). They are each divided further into four groups:

(i) W: clean (less than 5% finer than 0.075 mm); well graded (uniformity coefficient

C_u, greater than 4 for gravels or 6 for sands, and coefficient of curvature C_c between 1 and 3).

(ii) P: clean (less than 5% finer than 0.075 mm); poorly graded (C_u, less than 4 for gravels or 6 for sands, graded or gap-graded because C_c not between 1 and 3).

(iii) C: dirty (more than 12% finer than 0.075 mm); plastic clayey fines (PI greater than 7%, also plots above A-line in plasticity chart).

(iv) M: dirty (more than 12% finer than 0.075 mm); non-plastic silty fines (PI less than 4%, or plots below A-line in plasticity chart).

The soils are represented by symbols such as GW or SP. Border line materials are represented by a double symbol, as GW-GP.

The fine-grained soils are divided into three groups: inorganic silts (M), inorganic clays (C), and organic silts and clays (O). The soils are further divided into those having liquid limits lower than 50% (L), or higher (H). The distinction between the inorganic clays C and the inorganic silts M and organic soils O is made on the basis of a modified plasticity chart as mentioned above. Soils CH and CL are represented by points above the A-line, whereas soils OH, OL, and MH correspond to positions below. Soils ML, except for a few clayey fine sands, are also represented by points below the A-line. The organic soils O are distinguished from the inorganic soils M and C by their characteristic odor and dark color or, in doubtful instances, by the influence of oven-drying on the liquid limit. The unified soil classification system permits reliable classification on the basis of relatively few and inexpensive laboratory tests. It also provides a practicable basis for visual or field classification. Like all procedures based on grain size or the properties of remoulded materials, it cannot take into consideration the characteristics of the intact materials as found in nature. Hence, it can serve only as a starting point for the description of the engineering properties of soil masses or soil deposits.

2.6 Petrophysical properties

Electrical and magnetic properties of a material can be characterized by the static and dynamic behavior of electric and magnetic fields. These properties are electrical conductivity, dielectric permittivity, and magnetic permeability. The well known Maxwell equations, which relate the spacial and temporal variation of electric and magnetic field to each other, can be written in differential form as

$$\nabla \times \vec{E} = -\frac{\partial \vec{B}}{\partial t}, \tag{2.23}$$

$$\nabla \times \vec{H} = \vec{J} + \frac{\partial \vec{D}}{\partial t}, \tag{2.24}$$

$$\nabla \bullet \vec{D} = \delta, \tag{2.25}$$

$$\nabla \bullet \vec{B} = 0, \tag{2.26}$$

where $\vec{E}$ is the electric field vector (Vm^{-1}), $\vec{B}$ is the magnetic induction vector (Tm^{-2}), $\vec{H}$ is the magnetic field vector (Am^{-1}), $\vec{D}$ is the dielectric displacement current density vector (Cm^{-2}), $\vec{J}$ is the conduction current density vector (Am^{-2}), δ is the electric charge density (Cm^{-3}), and t is the time (s). The applicable constitutive equations, which combine the field quantities and the material properties, are

$$\vec{J} = \sigma \vec{E}, \tag{2.27}$$

$$\vec{D} = \epsilon \vec{E}, \tag{2.28}$$

$$\vec{B} = \mu \vec{H}, \tag{2.29}$$

where σ is the electrical conductivity (Sm^{-1}), ϵ is dielectric permittivity (Fm^{-1}), and μ is the magnetic permeability (Hm^{-1}).

2.6.1 Electrical conductivity

The interaction between soil and electromagnetic fields depends on soil particle size, mineral structure, mineral surface conditions and characteristics of pore fluid, as well as ion exchange capacity and properties of electrolytes. The ion movement direction in a soil-water system follows the direction of electric current. The influenced area is related to the magnitude of electrical charge and characteristics of the soil-water system.

The electrical conductivity σ is a quantity that characterizes electrical charge transport. It is an intrinsic property of any material. When an electric field $\vec{E}$ is applied, an electric current density $\vec{J}$ is established resulting from the displacement of various charged particles, such as electrons and or ions. The linear equation 2.28, an expression of Ohm's law, defines the electrical conductivity of an isotropic material. Alternatively, in applied geophysics the reciprocal value of conductivity, the electrical resistivity ρ is widely used,

$$\rho = \frac{1}{\sigma}. \tag{2.30}$$

Resistivity ρ has the SI unit of Ohm meter (Ωm) and the unit of conductivity is Siemens per meter (Sm^{-1}). Conduction current density $\vec{J}$ may be either electronic (the migration of loosely bound electrons in metals, sulfides) or electrolytic (the migration of ions in electrolyte). For granular soils, conductivity is mainly a function of electrolytic conductivity, which in turn is a function of void ratio, degree of saturation and the pore fluid parameters such as pressure, temperature and the content of Total Dissolved Solids (TDS). In the case of saturated granular soils, the bulk conductivity can be calculated using Archie's law as

$$\sigma = \frac{1}{F}\sigma_w, \tag{2.31}$$

where σ_w is electrolyte conductivity and F is the formation factor with

$$F = \frac{a}{\phi^m}, \tag{2.32}$$

where ϕ is porosity, m is the so called cementation factor and a is a further empirical parameter. These two empirical parameters characterize the soil texture controlled by the pore channel geometry of the soil (Schön, 1996). Meanwhile, for fine grained soils, electrical conduction becomes more complex since electrical double layers (EDL) are developed around the particles with negative surface charge. Significant electrical conductivity results from the presence of clay minerals (Olphen, 1977). The greater concentration of ions in the Stern and diffuse part of the EDL than in the bulk electrolytic solution gives rise to surface conduction and the diffusion of ions in these regions. The surface conduction and the diffusion of ions are the main reasons that volumetric mixing models to determine the electrical conductivity of composite material do not work for fine grained soils.

Complex conductivity

Complex conductivity is an expression that describes both conduction and charge storage processes. When an electric field is applied, it causes both an irreversible migration of electric charge and a reversible displacement of electric charge. The first process is characterized by the real part of conductivity of the soil $\sigma^{'}$ and the latter is characterized by charge separation or polarization of the soil that can be described by the the imaginary part of conductivity $\sigma^{''}$. The combination of both processes results in a complex conductivity as

$$\sigma^*(\omega) = \sigma^{'}(\omega) + i\sigma^{''}(\omega), \tag{2.33}$$

where i $= \sqrt{-1}$ is imaginary unit and ω the angular frequency ($\omega = 2\ \pi f$).
According to Börner and Schön (1991) and Weller and Börner (1996), in the low frequency range from 10^{-3} to 10^{3} Hz, the effective conductivity is a frequency-dependent complex parameter. The complex conductivity for a large variety of soils can be well expressed by the so-called constant phase angle model

$$\sigma^*(\omega) = (i\omega)^b \sigma. \tag{2.34}$$

As shown in equation 2.34, the separation of effective conductivity into real and imaginary parts yields

$$\sigma^{*\prime}(\omega) = (i\omega)^b \sigma^{'}, \tag{2.35}$$

$$\sigma^{*\prime\prime}(\omega) = (i\omega)^b \sigma^{''}, \tag{2.36}$$

where the exponent b with the value in range of 0.0001 to 0.03, characterizes the frequency dependence. The constant phase angle, which is defined as the ratio of imaginary part and real part of conductivity, is related to the frequency exponent b in equations 2.34 to 2.36:

$$\varphi = arctan\left(\frac{\sigma^{''}}{\sigma^{'}}\right) = \frac{\pi}{2} b. \tag{2.37}$$

Clays and other grain surface phenomena create an EDL with an interface conductivity as described in section 2.3. The real part of conductivity σ' contains both electrolytic conductivity determined by Archie's equation 2.31 and interface conductivity $\sigma_i^{'}$:

$$\sigma^{'} = \frac{\sigma_w}{F} + \sigma_i^{'}. \tag{2.38}$$

The imaginary part $\sigma_i^{''}$ results only from interface effects as

$$\sigma'' = \sigma_i^{''}. \tag{2.39}$$

In fine grained soils, the conductivity can be described by a real electrolytic volume conductivity and a complex surface conductivity as shown in equations 2.38 and 2.39 (Weller and Börner, 1996).

A sample holder was developed to measure complex conductivity of soil samples in laboratory. The conductivity and phase angle of soil samples in low frequency range from 3 mHz to 750 Hz were observed by Spectral Induced Polarization (SIP) method with a SIP-Fuchs central equipment. The measurement system consists of a sample holder (see Figure 2.14), a current source, a potential processing unit, a SIP-Fuchs unit and a notebook computer. A sinusoidal current signal of a single frequency is injected at the electrodes C1, C2 and potential difference at responding frequency will be observed at non-polarizing ring electrodes P1 and P2 made of silver chloride or platinum. Source and receiver signals are processed by the SIP-Fuchs central unit and conductivity magnitude and phase shift at all frequencies are recorded. The soil material was filled in a cylindrical sample holder with a dimension of 40 mm in diameter and 85 mm in length. The sample holder is made of a clear Polyvinyl Chloride (PVC) tube, which produces no spurious phase response. Two porous glass disks are positioned at either ends of the soil sample as filter. Two ring electrodes P1 and P2 with a distance of 55 mm were attached at the inner surface of the PVC tube. The soil samples were kept in an incubator at a constant temperature of 20 ^{0}C in order to reduce the influence of electromagnetic noise and temperature fluctuations. The measurements were repeated after 24 hours until the phase shift show a stable behavior.

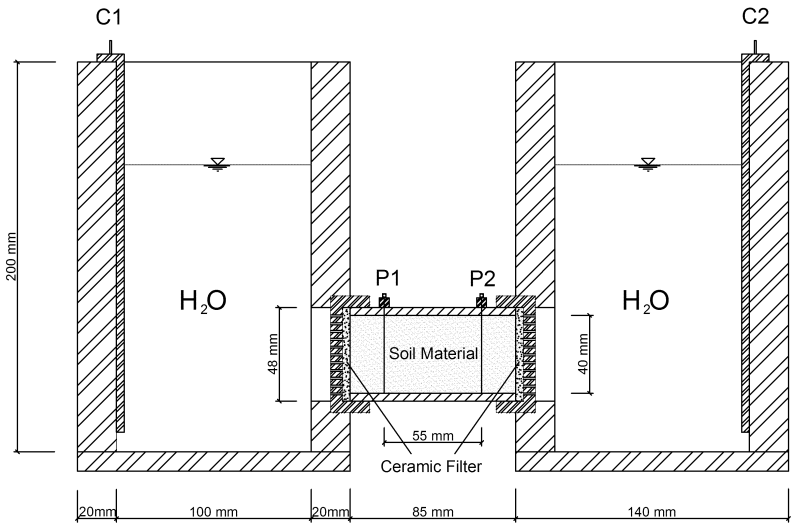


Figure 2.14: Sample holder for SIP measurement: A current signal is injected at electrodes C1 and C2. The corresponding potential is observed at electrodes P1 and P2.

2.6.2 Dielectric permittivity

When an electric field is applied to a material, in addition to a current of free charges, a local redistribution of bound charges to new equilibrium positions occurs. This phenomenon of charge redistribution is called polarization which results in an induced field and tends to oppose the applied field. The dielectric permittivity is defined as a material's ability to maintain charge separation or polarization. It is a measure of the capacity of a material to reduce the strength of an electric field. In practice, the dielectric permittivity of materials is normalized by the dielectric permittivity of free spaces (ϵ_0) and referred to as relative dielectric permittivity or dielectric constant ϵ_r:

$$\epsilon_r = \frac{\epsilon}{\epsilon_0}, \tag{2.40}$$

with $\epsilon_0 = 8.854{\times}10^{-12}$ Fm^{-1}.

In general, like electric conductivity, dielectric permittivity is a complex quantity. There are two cases to consider when a dielectric is subject to an alternating electric

field, depending on the frequency of the applied field, the temperature and other parameters. In the first case, there would be no measurable phase difference between a displacement current density and an applied electric field indicating that the polarization is in phase with the applied field. In the second case, there would be a noticeable phase difference between $\vec{D}$ and $\vec{E}$, and then relative permittivity can no longer be considered to be a constant (Canan, 1999). The real part of the dielectric permittivity is a measure of how much energy from an external electrical field is stored in a material. The imaginary part of permittivity is related to electrical conductivity. The ratio between imaging and real part of permittivity is a measure of dissipation or loss of energy of an external field in a material (Balanis, 1989). The effective dielectric permittivity is defined as

$$\epsilon^* = \epsilon^{'} + i\epsilon^{''}. \tag{2.41}$$

The ratio of the imaginary and real part gives the loss tangent δ as

$$\delta = arctan\left(\frac{\epsilon^{''}}{\epsilon^{'}}\right). \tag{2.42}$$

The relation between electrical conductivity and dielectric permittivity is expressed as

$$\epsilon^* = \epsilon + i\frac{\sigma}{\omega} \tag{2.43}$$

and

$$\delta = \varphi + \frac{\pi}{2}. \tag{2.44}$$

As mentioned in equation 2.40, in practice, the complex dielectric constant (ϵ^*) is equivalent to relative permittivity and it is expressed as

$$\epsilon_r = \epsilon_r^{'} + i\epsilon_r^{''}. \tag{2.45}$$

In the laboratory, the complex dielectric constant of soil sample was measured with an HP RF Network Analyzer in the frequency range from 1 MHz to 1 GHz. The Network Analyzer consists of a signal generator, a sample holder and a computer. The sample holder may be regarded as a capacitor. The source signal, consisting of a frequency scan of discrete frequencies, is transmitted through the sample holder. From the measured response, the magnitude and phase data are calculated at discrete frequencies. In preparation of sample, the soil material is filled in a Teflon or Perspex sample holder ring with a diameter of 13 mm and a thickness of 4 mm. The filled ring is capped by aluminum foil to guaranty good contact to the electrodes and positioned between the two electrodes of the capacitor. The measurements were repeated with different water content of soil sample. The influence of the ring is eliminated in a special correction program.

2.6.3 Magnetic permeability

Magnetic properties describe the behavior of material under the influence of an external magnetic field. Magnetic properties in soils are largely a consequence of the presence of different forms of iron or iron minerals. The magnetic permeability of a material μ is defined in equation 2.29.

The physical background for the existence of magnetic behavior in minerals is the magnetic moment produced by electrons orbiting their nucleus and spinning around their axis. In many types of material, the overall magnetic moment is zero because the orbital and spin components even out. When a mineral with zero magnetic moment is placed in a magnetic field the electron motions will rearrange so that a net magnetic moment is in the direction opposite to the applied field. These types of minerals are called *diamagnetic*. In contrast, when minerals with a small net magnetic moment get subjected to a magnetic field the magnetic spins will attempt to line up in the direction of the magnetic field. These types of minerals are called

paramagnetic. In some minerals, the interaction between electron spin and orbital movement in adjacent atoms causes these minerals to behave as active magnets. These types of minerals are called *ferromagnetic* when all magnetic moments line up in the same direction, or *ferrimagnetic*, when a certain part of the magnetic moments line up in the opposite direction. A special group of minerals are those in which the electron interaction leads to magnetic moments being aligned in opposite directions. These minerals with a net magnetic moment of zero are called *antiferromagnetic*. Many books and review papers have addressed the physical background of magnetic minerals in general such as Guéguen and Palciauskas (1994), Schön (1996), Evans and Heller (2003) and magnetic soils in particular Mullins (1977), Dearing et al. (1996) and Van Dam et al. (2004).

Table 2.2: Mass magnetic susceptibility of some common minerals in soil (after Mullins, 1977; Schön, 1996 and Potter et al., 2004).

Magnetic type	Mineral	Mass magnetic susceptibility (10^{-8} m^3 kg^{-1})
Diamagnetic		
	Quartz	-0.55
	Water	-0.9
	Calcite	-0.3 to -1.4
	Kaolinite	-1.9
	Feldspar	-0.49 to - 0.67
Paramagnetic		
	Illite	15
	Chlorite	13.6 to 52.5
	Pyrite	2
	Montmorillonite	2.7
	Bentonite	5.8
Ferrimagnetic		
	Magnetite	1^4 to 11^4

In natural soil diamagnetic, paramagnetic and ferrimagnetic minerals are among the most common constituents. The mass magnetic susceptibility of some common minerals naturally occurring in soil are compiled in Table 2.2.

In performing magnetic measurements, the sum of all magnetic moments or the total magnetic moment per unit volume is regarded. This is called the magnetization M

Am^{-1}. The ratio of the magnetization to the magnetic field H, which is inducing it, is called the volumetric magnetic susceptibility χ_v,

$$\chi_v = \frac{M}{H}. \tag{2.46}$$

The magnetic permeability μ relates to volumetric magnetic susceptibility χ_v by the equation

$$\mu = \mu_0(1 + \chi_v), \tag{2.47}$$

where $\mu_0 = 4\pi\times10^{-7}$ ($VsA^{-1}m^{-1}$) is the magnetic permeability for vacuum, and

$$\mu_r = 1 + \chi_v \tag{2.48}$$

is called the relative magnetic permeability. Both volumetric magnetic susceptibility χ_v and relative magnetic permeability μ_r are dimensionless quantities in the International System of unit (SI).

The total magnetic moment per unit mass divided by the field is called the mass normalized susceptibility χ_m. It is related to χ_v via the density of the material ρ as

$$\chi_m = \frac{\chi_v}{\rho} \tag{2.49}$$

and is given in the unit m^3kg^{-1}.

There is a wide variety of methods available for the measurement of magnetic susceptibility of which only a few are appropriate to soil studies. The simplest and most reliable method of measuring the susceptibility of soils is the alternating field bridge design (Mullins, 1977). In this study, all magnetic susceptibility of soil sample were measured by Kappabridge KLY-2 equipment with an operating frequency of 920 Hz. The equipment comprises measuring unit and a standard pick-up unit. Its operation is based on measurement of inductivity changes in a coil due to the soil

sample. By inserting dry soil sample into pick-up unit its inductivity is changed. The induced signal of unbalance will be compensated immediately by the effect of the feedback loop. A voltage proportional to the inductivity change and thus to the susceptibility measured will appear on the display. The measurement was repeated three times and the average value will be used.

Chapter 3

Description of site locations and soil properties

3.1 Introduction

The soil samples from the Red River Delta in the North of Vietnam were investigated in this study. Geotechnical and petrophysical properties of these samples were determined in laboratory. Another set of soils extracted from dike body and its foundation in Sachsen and Sachsen Anhalt, Germany was also investigated. The origins of soil samples from investigation sites in Vietnam and Germany will be briefly described in this chapter. To describe and explore the geotechnical and petrophysical properties of all soil samples from different sites, a tool of descriptive statistics is required. Several measures exists, but a few are used as basic concepts of probability theory to describe briefly the uncertainty of a specific variable. These measures are widely used in an attempt to present the central tendency and dispersion behaviors of a certain data set. A brief summary on basic descriptive statistic is given. Details and further descriptive tools of data can be found in many statical analysis textbooks. In the field of geosciences, the descriptive tools are well documented in Davis (2002)

and Baecher and Christian (2003).

The common measures of central tendency are the mean and median. The *mean* is the arithmetic average of a data set. The *median* is the middle value of a set of data when the observations are ranked from smallest to largest. Given a variable $x_1, x_2, x_3, ..., x_n$, n is the number of sample in a given data set, the mean denoted $\overline{x}$ can be seen as the center of gravity of the frequency distribution along the variable axis,

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i. \tag{3.1}$$

The common measures of dispersion or spread are variance, standard deviation, range and inner quartile of frequency distribution. The *variance*, s^2, is the moment of inertia of frequency distribution around the mean and can be determined as

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2. \tag{3.2}$$

The variance has an unit of variable squared, it occasionally makes inconvenience in comparison. The square root of variance has the same unit as the original variable. It is named *standard deviation* and is defined as

$$s = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2}. \tag{3.3}$$

To measure the mutual variability of two variables, the joint variation of two variables, which is called *covariance*, is widely used. The covariance of two variables x_j and x_k can be calculated as

$$COV_{jk} = \frac{1}{n-1}\sum_{i=1}^{n}(x_{ij} - \overline{x}_j)(x_{ik} - \overline{x}_k). \tag{3.4}$$

In order to estimate the degree of interrelation between two variables, the *correlation coefficient* r is used. The correlation is the ratio of the covariance of two variables to the product of their standard deviations as

$$r_{jk} = \frac{COV_{jk}}{s_j s_k}. \tag{3.5}$$

The *range* is the difference between the largest and the smallest values. It has poor statistical properties in the sense that it is sensitive to extreme values in a data set, however it is easily evaluated and therefore often useful in a quick estimate.

The *inner quartiles* of a set of data, denoted the lower quartile (25%) and the upper quartile (75%), are the data values for which one-quarter of the data are smaller and one-quarter larger, respectively in frequency distribution.

In comparison of dispersion among different dimensional variables, the quantity *coefficient of variation* (CV) is usually used. It is defined as the standard deviation divided by the mean

$$CV = \frac{s}{\overline{x}} \tag{3.6}$$

which expresses relative dispersion as a fraction.

All estimates are subjected to errors. Sample information is never complete and its uncertainty should be measured. The solution is based on the standard deviation of the means or standard error (SE) in short. The standard error is defined as

$$SE = \frac{s}{\sqrt{n}}. \tag{3.7}$$

It follows from this equation, that the larger the number of samples, the smaller the likely standard error.

3.2 Brief description of site locations and soil sampling in Vietnam

Soil samples were collected on the hydraulic works on the dike or civil engineering construction sites close to the dike system. In total 120 soil samples at different depth were collected at seven site locations in Thai Binh and Nam Dinh provinces in Vietnam whose dike system is the densest and among the most highly endangered to inundation and flood in Vietnam. Some of the undisturbed samples were used for the determination of geotechnical parameters in Vietnam and almost disturbed samples were used to investigate petrophysical properties in Germany.

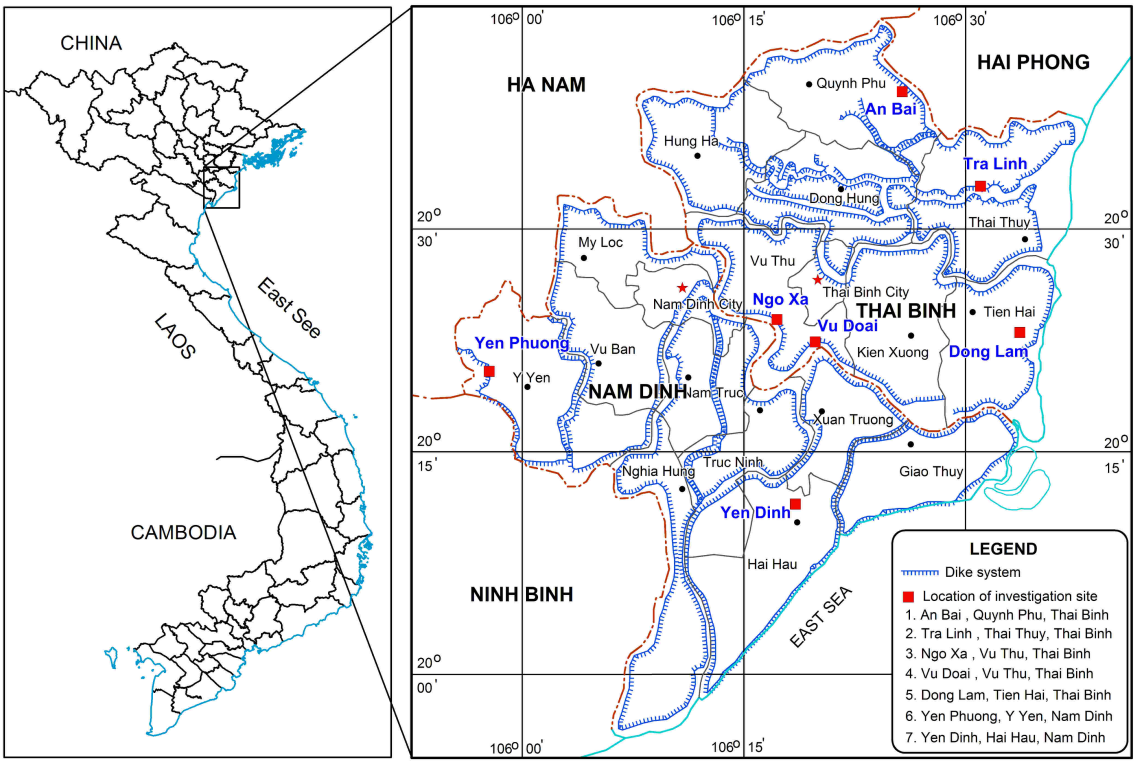


Figure 3.1: the dike system in Nam Dinh and Thai Binh provinces and site locations (extracted from the map of dike system in the northern Vietnam, The Dike Management and Flood Control, Hanoi, Vietnam).

3.2.1 An Bai, Quynh Phu, Thai Binh province

This is a construction site of an industrial area at An Bai town situated at $20^0 40'$ North and $106^0 25'$ East close to Hoa river dike as marked in Figure 3.1. Fifty mainly clay and sandy clay soil samples were extracted from five boreholes at different depth. The deepest sample reaches the depth of 35 m under the surface. All the soil samples were divided, partly for geotechnical experiments and the rest for petrophysical analysis.

3.2.2 Tra Linh, Thai Thuy, Thai Binh province

On upgrading the tidal sluice of Tra Linh on Diem Ho river dike, a geotechnical investigation was performed. Site location is situated at $20^0 34'$ North and $106^0 31'$ East as shown in Figure 3.1. Eleven soil samples were extracted at three boreholes BH1, BH2 and BH3. One sample was obtained at BH1, four samples at BH2 and five samples at BH3. The deepest sample was taken at 40 m under the surface in BH3. All samples were divided apart, undisturbed samples for geotechnical test and the rest for petrophysical analysis in laboratory.

3.2.3 Dong Lam, Tien Hai, Thai Binh province

A new pottery factory was planed to build up here and a geotechnical investigation was performed to evaluate its foundation. The location, located at $20^0 23'$ North and $106^0 33'$ East as pointed in Figure 3.1, is approximately in 1 km distance to the sea dike profile. The soil samples from this site were extracted from two boreholes 75 m apart in distance. Seven sample were extracted from a borehole with depth of 27 m. From the other borehole reaching a depth of 22 m, four samples were extracted. Eleven soils samples were collected of which one sample is sand, three silt samples and seven clay samples. All samples were transported to laboratory for geotechnical

and petrophysical analysis.

3.2.4 The dike monitoring system at Ngo Xa

The weak dike segment at Ngo Xa belonging to the Red river dike system in Thai Binh province located at $20^0 24'$ North and $106^0 17'$ East was selected to install a dike monitoring system. Before installation, four boreholes BH1, BH2, BH2A and BH3 were completed for soil sampling along a cross-section of the dike as shown in Figure 3.2. Borehole BH1 reaches a depth of 10,2 m and borehole BH2 at depth of 8,0 m. Five samples were taken at BH1, three samples at BH2, five samples at BH2A and 3 samples at a water table monitoring well in distance of 50 m from the site. A total of 18 samples were taken for analyzing geotechnical properties in laboratory. Other eight samples were collected for petrophysical analysis. A set of six samples was extracted at different depth at BH1 and two other were taken at BH2 and BH3.

A permanent geoelectrical system consisting of 50 electrodes with a spacing of 0.8

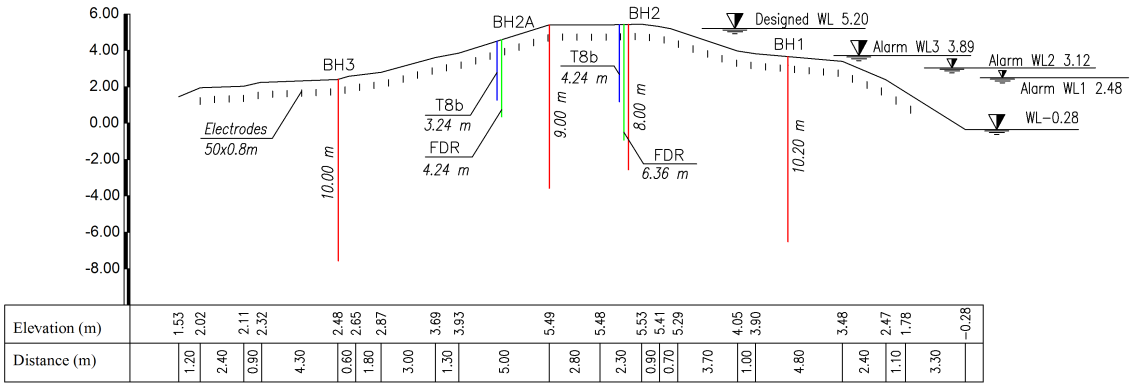


Figure 3.2: Dike monitoring system at Ngo Xa, Vu Thu, Thai Binh province.

m was installed across dike section at a depth of 0.5 m to monitor the behaviour of subsurface as a function of time. Two EnviroSCAN's probes of sensors were penetrated in subsurface to depths of 6.36 m and 4.24 m to monitor the water content variation during time. The EnviroSCAN sensor technology utilizes Frequency Domain Reflectometry (FDR) to measure soil water content. The changes of pressure

and temperature were observed by two tension meters installed in a depth 4.24 m and 3.24 m below the surface. The positions and depths of penetration of all boreholes and installed equipments are shown in Figure 3.2. The details of the dike monitoring at Ngo Xa site are described in Weller et al. (2010).

3.2.5 Vu Doai, Vu Thu, Thai Binh province

A new intake sluice for irrigation purpose was planed to be build on the dike of Vu Doai belonging to Tra Ly river dike. The location of the site is $20^0 23^{'}$ North and $106^0 19^{'}$ East as pointed in Figure 3.1. Twelve soil samples of which four samples and eight samples were extracted from two boreholes BH1 and BH2, respectively. There are four sand samples, four silt samples and the rest are clay. All samples were prepared for both geotechnical tests and petrophysical investigation in laboratory.

3.2.6 Yen Dinh, Hai Hau, Nam Dinh province

This is a civil engineering work for military purpose of Hai Hau district, Nam Dinh province, situated at $20^0 12^{'}$ North and $106^0 18^{'}$ East, approximately 5 km away from the dike as shown in Figure 3.1. Four soil samples were obtained from borehole BH1 and five samples from borehole BH2. Five of the samples are sand and the rest four sample are clay. All samples were divided apart, one for geotechnical analysis and the other for petrophysical analysis in laboratory.

3.2.7 Yen Phuong, Y Yen, Nam Dinh province

This is a very weak dike segment at Yen Phuong on the left Day river dike. It is located at $20^0 19^{'}$ North and $105^0 58^{'}$ East, as shown in Figure 3.1. Seventeen soil samples were collected which three sandy clay samples from one borehole and fourteen clay samples were taken from the slope of the dike body at the depth of

50 cm below the surface. All the samples from this site were analyzed only for petrophysical properties in laboratory.

3.3 Brief description of site locations and sampling in Germany

On the frame work of the project titled "Systematische Evaluierung exsistierender und innovativer Methoden zur Schwachstellenanalyse und Strukturerkundung von Deichen" (DEISTRUKT) funded by Ministry of Education and Research, Germany, two dike locations on Mulde river dike system were chosen to investigate. Petrophysical properties of dike material at two locations namely Friedersdorf and Loebnitz and one another location of Polder Roesa as shown in Figure 3.3 were investigated in laboratory.

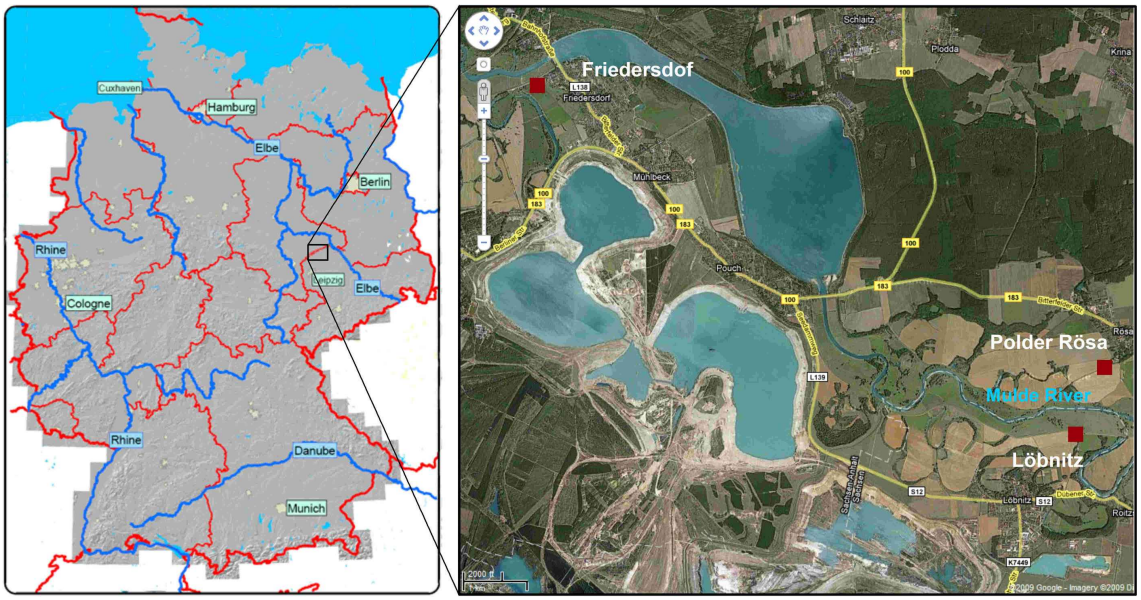


Figure 3.3: River system in Germany and site locations (extracted from Google Map).

Two boreholes QP 1/5 at the station 2+345 (new dike) and QP 4/5 at the station 2+760 (old dike) on crest of dike were sampled in Loebnitz, seven and five soil

samples were extracted along two boreholes, respectively. Seven other samples were collected at the flank and foot of two dike cross sections. Most samples are clay, some are fine sand and silt. At cross section station 2+345, an electrical tomography was performed to observe the in-place resistivity distribution in the dike body.

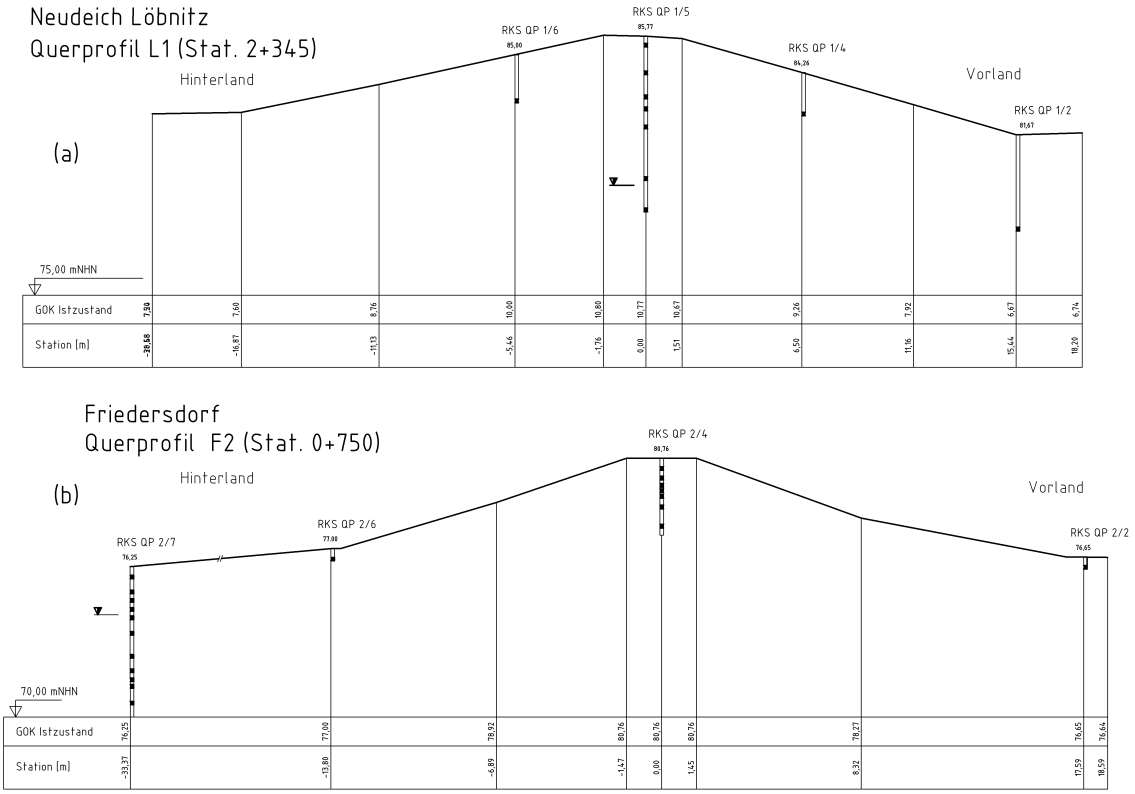


Figure 3.4: Soil sampling at the dike of Germany: (a) Soil sample locations at the dike of Loebnitz at station 2+345; (b) Soil sample locations at the Friedersdorf dike at station 0+750.

In the Friedersdorf dike, soil sampling was conducted at the station 0+750 cross section. Seven samples along borehole QP 2/7 at the crest of the dike and eleven samples along borehole QP 2/4, 10 m apart from dike foot were extracted. Two other sample were taken at two boreholes QP 2/2 and QP 2/6 at the toes of the dike.

Opposite to Loebnitz dike through Mulde river, the dike Polder Roesa was also

investigated. Fifty soil samples from thirty boreholes along the dike segment from station 0+300 to station 6+100 were collected. The majority of soils are sand and clay. The samples were taken from the surface to a depth of about six meters from boreholes at the dike crest, the dike slope and the dike toe.

3.4 Geotechnical properties of soils

More than a hundred soil samples were collected and tested in laboratory for analyzing geotechnical properties of soils from six locations in Nam Dinh and Thai Binh provinces in northern part of Vietnam. The properties of soil samples including grain size distribution, Atterberg limits, water content, porosity, density, cohesion and angle of internal friction are statistically described in the appendix.

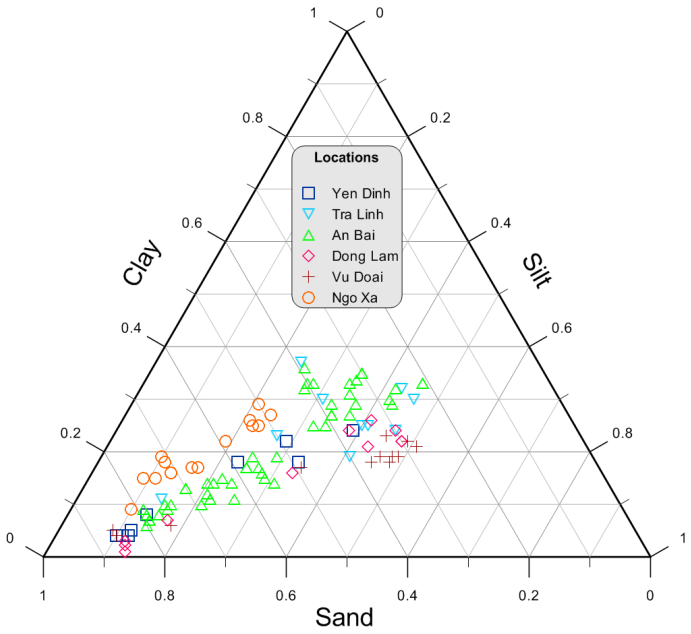


Figure 3.5: Grain size distribution of soils from six investigation sites in Vietnam.

Soil grain size ranges are plotted in a ternary graph as shown in Figure 3.5. The sand content varies from 21% percent to 87%, while clay fraction ranges from 1% to

37%. The silt fraction is in the range of 9% to 51%. In general, the grain size of soil samples from all locations is mainly distributed from the leftmost corner, which is presented by clean sand samples, to the center of the triangle where the percentages of clay, silt and sand fractions become similar.

A plasticity chart is widely used to classify fine-grained soils in accordance with

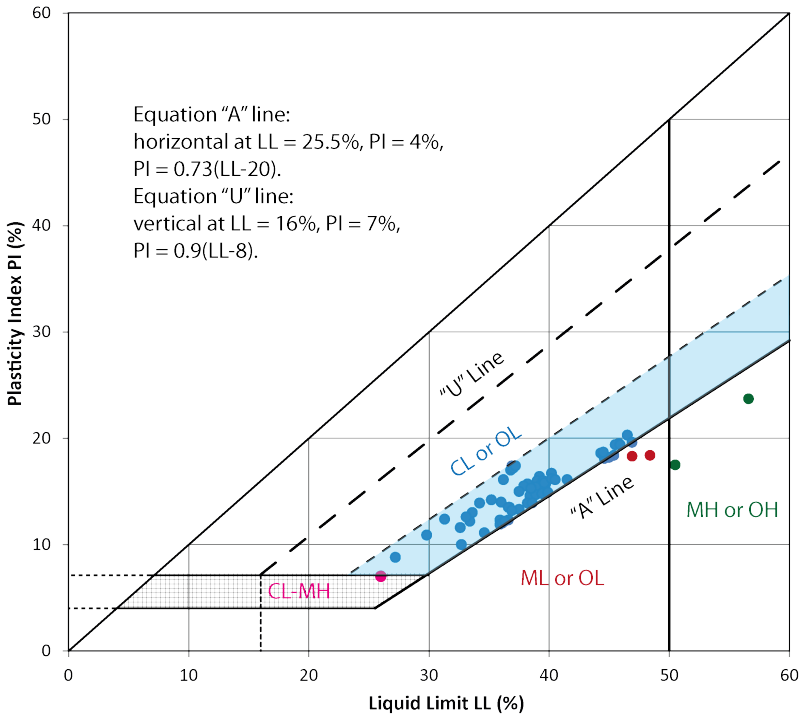


Figure 3.6: Classification of fine-grained soils from Vietnam in ASTM standards. N.B.: CL: Lean clay; ML: Low plasticity silt; MH: High plasticity silt; OL: Organic silt; OH: Organic clay.

ASTM standard. Almost all fine-grained soils from Vietnam as shown in Figure 3.6 fall in the area between "U" line and "A" line. According to ASTM D2487 (2000), the soil within this area is classified as lean clay (CL) which is of low to medium plasticity because it contains a large proportion of silt or sand. Furthermore, these soils are in a even smaller area (in blue color), where the soil is composed predominately by illite mineral Holtz and Kovacs (1981) and Mitchell (1993). A pair of samples is characterized as organic clay (OH) with high liquid limit and plasticity index and another one as low plasticity silt (ML).

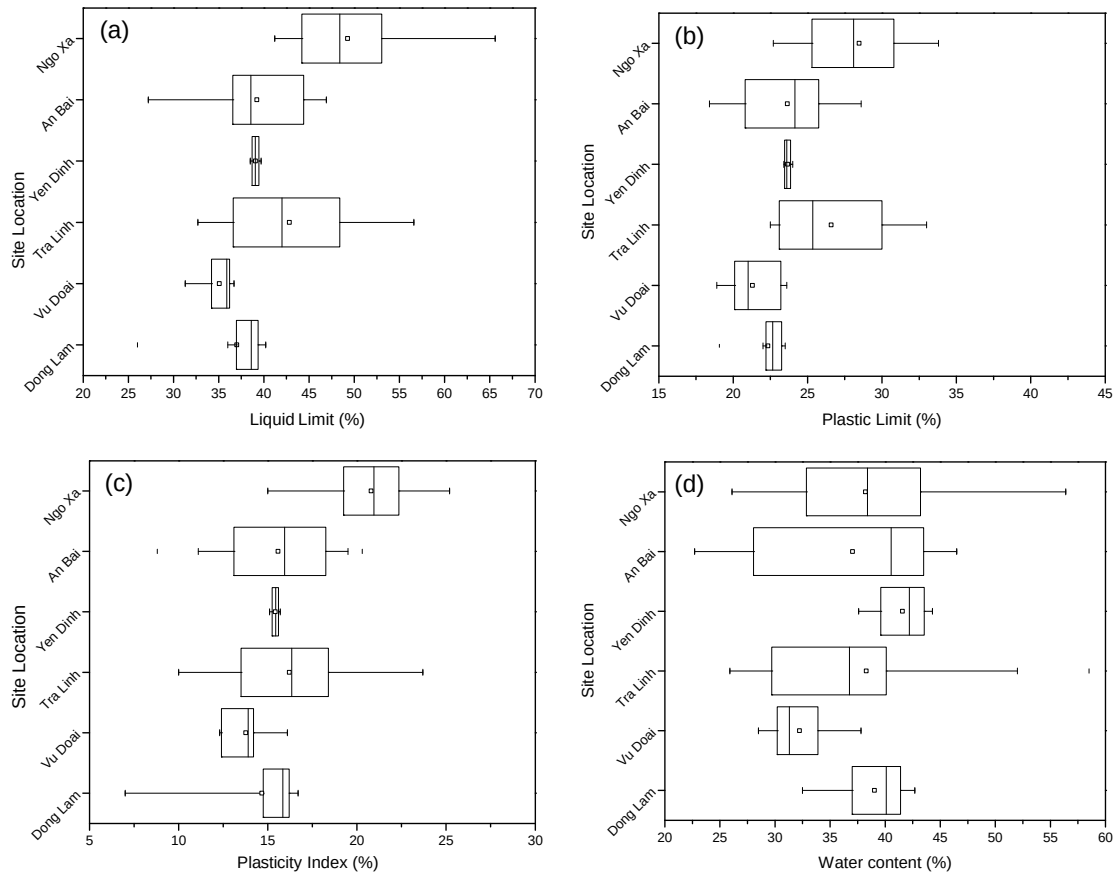


Figure 3.7: Distribution of Atterberg limits and water content of soils: (a) Liquid limit; (b) Plasticity limit; (c) Plasticity index; and (d) Water content.

The distributions of Atterberg limits and water content of the soils are presented as box-and-whisker plots in Figure 3.7. It is a convenient way to statistically and graphically describe groups of data set through five summarizing numbers: the minimum, lower quartile, median, upper quartile, and the maximum. A boxplot may also indicate which observations, if any, might be considered as outliers. At three locations of Dong Lam, Vu Doai and Yen Dinh, the values are in smaller and narrower ranges compared to the other locations. At Dong Lam location, all soil samples are lean clay, except one sample is low plasticity silt (ML) with low liquid and plastic limits of 26% and 19%, respectively. This sample is indicated as an outlier as shown in Figure 3.7(a) and Figure 3.7(b). Compared to the other sites, Vu

Doai has the lowest range of water content that results in the least values of liquid limit and plastic limit range and plasticity index as well.

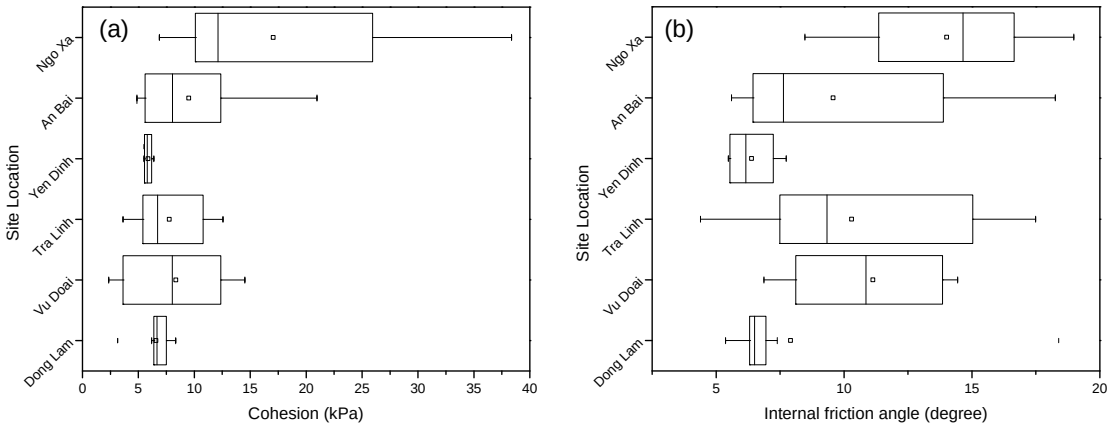


Figure 3.8: Distributions of cohesion and internal friction angle of soils: (a) Cohesion; (b) Internal friction angle.

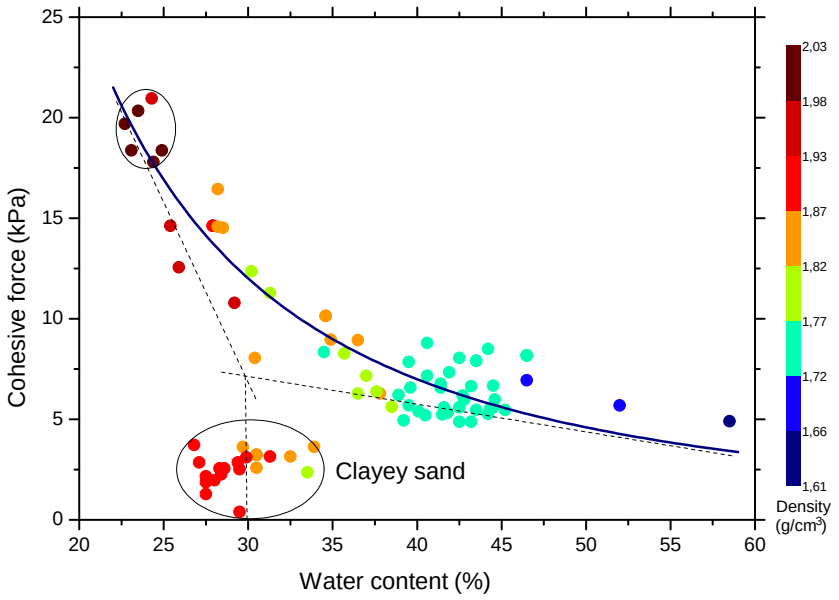


Figure 3.9: Cohesive force versus water content of soils with a reference to density

The distribution of strength parameters, cohesion and internal friction angle of cohesive soils, is presented in Figure 3.8. At Ngo Xa site, the soil strength is characterized by a higher value range compared to the other sites. In fact, the soil

samples from this site were mainly collected at the dike body, while at other sites, the soil mainly originated from dike foundation. At Dong Lam site, the silt sample (ML) as mentioned above is also presented as an outlier.

On the Mohr-Coulomb failure criterion, the shear strength of soils is measured by two components, cohesion force (c) and internal friction angle (ϕ). The higher the values of c and ϕ, the higher the shear strength of soils. The shear strength of soils is essential for any kind of an earth work's stability analysis. It is therefore important to determine the shear strength parameters as reliable as possible. The time-consuming triaxial tests are most appropriate in this case. But in fact, direct shear tests are mostly applied to determine the shear strength of soils as described in chapter 2. Thermann et al. (2009) described the shear strength parameters of soils determined by the direct shear tests and explained the influencing factors and their significances.

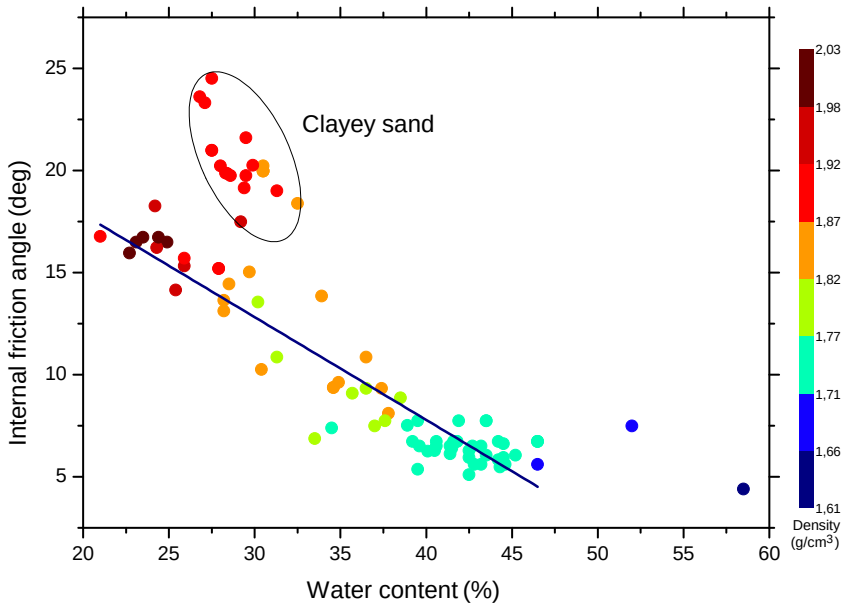


Figure 3.10: Relation between water content and internal friction angle with a reference to density.

The engineering behavior of a soil depends on compositional and environmental

factors (Mitchell and Soga, 2005). Among those water content effects significantly the shear strength of the soils. Figure 3.9 and 3.10 show the variation of internal friction angle and cohesive force responding to water content of soils. Cohesion within a soil mass is due to a variety of inter-particle forces. As water content increases, cohesion decreases. An increase of water content causes greater separation of clay particles and softening of soil cements. As shown in Figure 3.9 and 3.10, the clayey sand samples with moderate density, lower cohesion but higher angle of internal friction appear separately as illustrated by an encircled ellipse. The cohesion decreases exponentially with the increase of water content as $c = 7143w^{-1.88}$ and R^2=0.87, while the internal friction angle decreases linearly as $\varphi = 27.90 - 0.5w$ with $R^2 = 0.89$. Six clay samples of low water content, high density and high cohesion, presented in the top left of Figure 3.9 are stiff clays at a depth of greater than 25 m that might, in term of hydrogeology, belong to impervious aquitard layer (Do, 1996). Two organic clay samples (OH) with lowest density and highest water content, above 50%, seem to be outliers among clay soils (CL) as shown in Figure 3.10. In an investigation of cohesive unsaturated soil's strength behavior, Kato et al. (2000) presented a decrease of cohesion with increasing of water content and suggested that at a lower water content, less than the liquid limit of the soil in question, the cohesion decreases steeper than that of higher water content. In this study, the experimental result shows the decrease of cohesion approaching two tangents crossing at a water content of 30%. Al-Shayea (2001) presented various aspects of the behavior of synthetic clay-sand mixtures. A linear decrease of shear strength with an increase of water content has been observed at water contents lower than 25%. In this study, shear strength of various undisturbed soils were investigated with the water content varying from 20% to 60%. The soils with clay fraction lower than 12%, can be classified as clayey sand or sand with clay, scattered into group in red and orange color as shown in Figure 3.11.

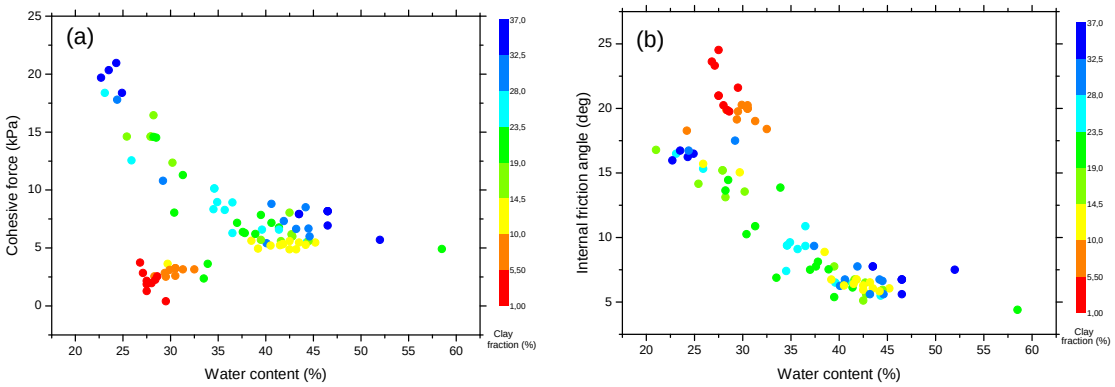


Figure 3.11: Soil shear strength behavior with various water contents with a reference to clay fraction: (a) Relation of water content and cohesion; (b) Relation of water content and internal friction angle.

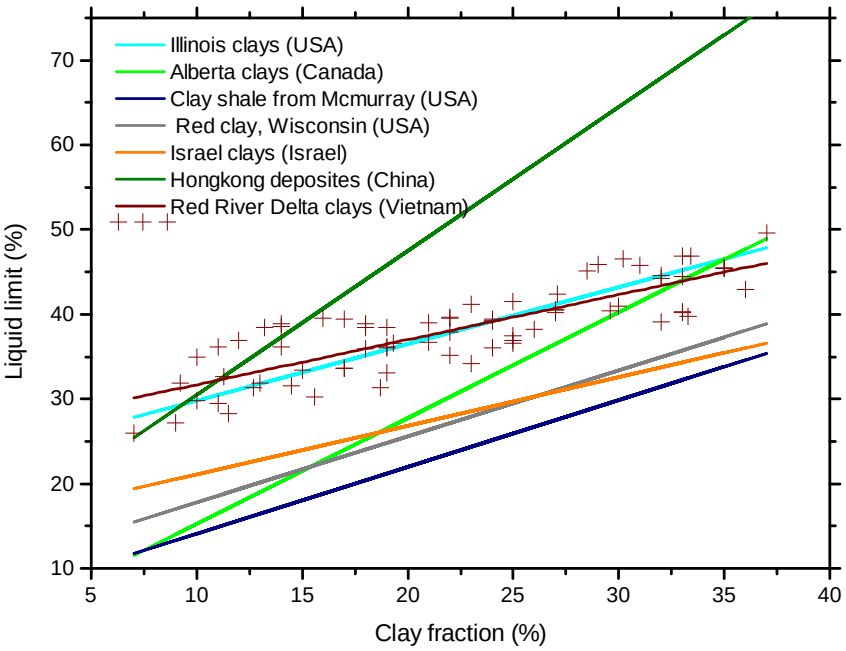


Figure 3.12: Liquid limit versus clay fraction.

Many investigations have presented on the relation of clay fraction and liquid limit from different type of clay soils as well as sedimentation environment conditions such as Odell et al. (1960), Brown and Mengel (1983), Smith et al. (1985), Yin (1999). In Figure 3.12, the relation between clay fraction and liquid limit of soils in the Red River Delta of Vietnam is plotted in comparison with those from other

Table 3.1: Linear regression of clay fraction and liquid limit with α being the slope and β the intercept.

Origin of soil samples	Coefficients		References
	α	β	
Illinois clays, USA	0.67	23.10	Odell et al. (1960)
Alberta clays, Canada	1.25	2.75	Hamilton (1966)
Clay shale, Mcmurray, USA	0.19	6.19	Dusseault and Scafe (1979)
Red clays, Wisconsin, USA	0.78	10.00	Brown and Mengel (1983)
Israel clays, Israel	0.58	15.35	Smith et al. (1985)
Hongkong deposits, China	1.70	13.50	Yin (1999)
Red River Delta clays, Vietnam	0.53	26.37	New data

locations found in the literature. These linear relations are tabulated in Table 3.1 where α and β are slope and intercept, respectively. The soils from Vietnam as described comprise predominantly illite as clay mineral. They are sedimented in alluvial or marine environments. The soils from Illinois (the name originating from its type of clay, illite) are obviously dominated by illite and deposited in alluvial environment. This results in almost coincident lines of soils from the Red River Delta and Illinois as seen in Figure 3.12. The Hongkong soils containing a great amount of montmorillonite appears apart from the others.

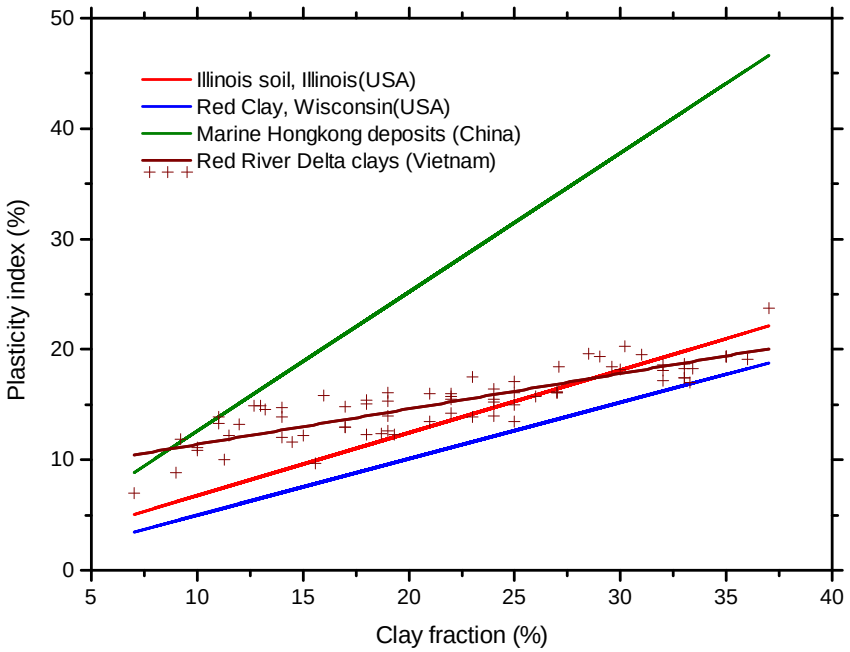


Figure 3.13: Plastic index versus clay fraction.

Table 3.2: Linear regression of clay fraction and plastic index with α being the slope and β the intercept.

Origin of soil samples	Coefficients		References
	α	β	
Illinois clays, USA	0.57	1.09	Odell et al. (1960)
Red clays, Wisconsin, USA	0.51	-0.10	Brown and Mengel (1983)
Hongkong deposits, China	1.26	0.00	Yin (1999)
Red River Delta clays, Vietnam	0.32	8.20	New data

Figure 3.13 shows the linear relations of clay fraction and plasticity index with linear regression coefficients tabulated in Table 3.2. The soils from Vietnam appear closely to Illinois soils while the marine Hongkong deposites exhibit a steeper slope because of the presence of montmorillonite mineral in the soil.

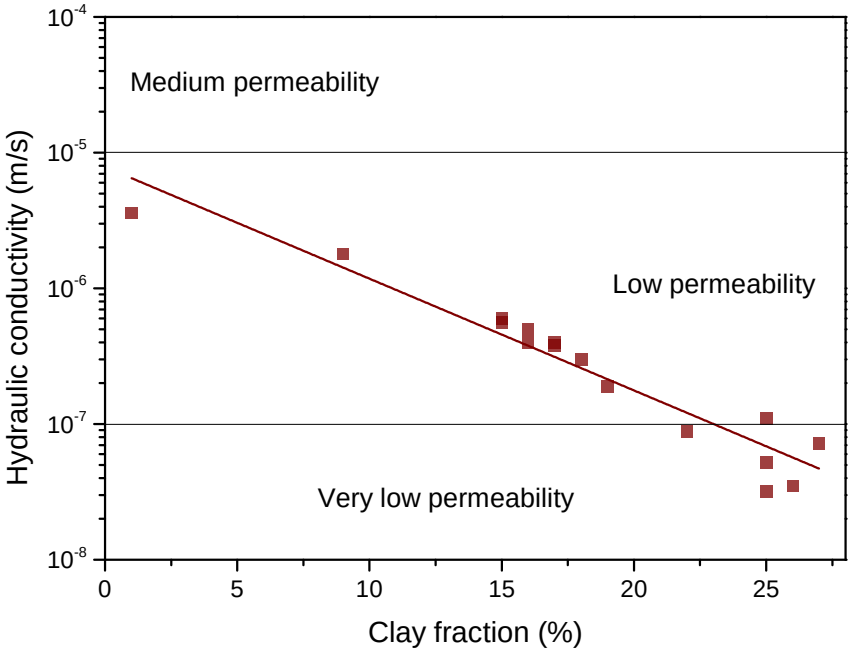


Figure 3.14: Hydraulic conductivity versus clay fraction.

Figure 3.14 shows a relation between clay fraction (CF) and hydraulic conductivity (k) of 16 soil samples from the dike monitoring system at Thai Binh, Vietnam. All the soil samples are of low to very low permeability. Hydraulic conductivity decreases by the increase of clay fraction as $Log(k) = -0.08 \times CF - 5.10$ with $R^2 =$

0.92. At the intercept, $k = 7.94 \times 10^{-6}$ m/s, the sand sample is located in the dike foundation at the depth of 9.2 m. The other soil samples are clays located in the dike body with a lower hydraulic conductivity.

3.5 Clay mineralogy of soils

Clay mineralogy of soils was analyzed by the widely used X-Ray diffraction method (XRD). The mineral crystal structure of soils is identified by known wavelength of X-Rays passing through the soil sample. The X-Rays are diffracted by the lattice of the crystal to give a specific pattern of peaks at different angle and intensity. The angle of incidence between glass of soil sample and X-ray beam θ at the peaks were measured and the lattice spacing d were calculated using Bragg's law as described in chapter 2. The identification of complex mixtures of minerals requires experience and knowledge of the relative intensity of different peaks. The crystal structure of soils from Vietnam was identified as illite, kaolinite, and chlorite as seen in Figure 3.15 where the abscissa shows the value of 2θ in degrees and the ordinate represents the intensity of the reflection.

The soils from Germany comprise kaolinite, illite, chlorite, and a small amount of smectite. Kaolinite shows a strong predominance over the others. The soils from Vietnam have illite mineral as dominant constituent and the rest is a smaller amount of chlolite and kaolinite. As shown in Figure 3.16, the clay mineral contents of two typical clayey soils representing the soils from Germany and Vietnam exhibits the dominant clay mineral as kaolinite (51%) and illite (60%), respectively. Table 3.3 shows the clay mineral contents of soils from various locations of Vietnam and Germany, resulting from quantitative clay mineralogical analysis.

According to Holtz and Kovacs (1981) and Mitchell (1993), the classification of

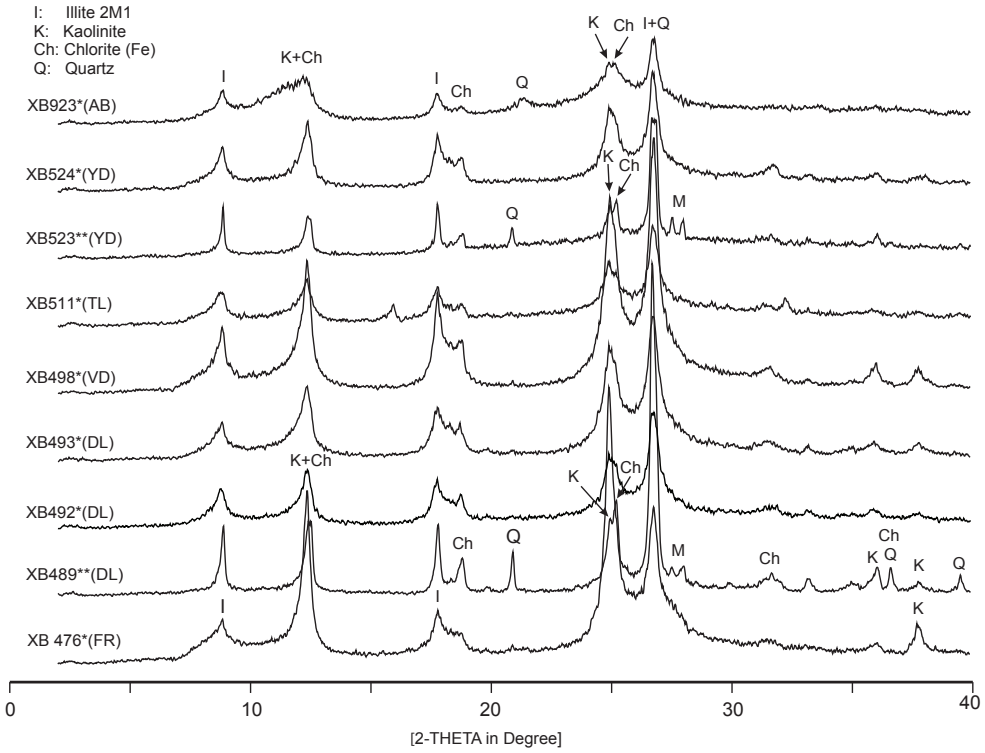


Figure 3.15: Clay mineralogical analysis of oriented soil samples: (*) soils with grain size $< 2\ \mu m$; (**) soils with grain size $< 5\ \mu m$. Site locations: AB: An Bai; YD: Yen Dinh; TL: Tra Linh; VD: Vu Doai; DL: Dong Lam, Vietnam; and FR: Friedersdorf, Germany.

soils is related to the type of clay minerals as illustrated in the palsticity chart of Figure 2.13. The fine-grained soils from Vietnam fall in the area of illite mineral in the plasticity chart as seen in Figure 3.6, which agrees with the fact that all soils from Vietnam in Table 3.3 having illite as a dominant clay mineral.

The specific surface area (SSA) was measured for ten fine-grained soil samples from Vietnam with different clay fractions. Clay fraction in the soils and the SSA are compiled in Table 3.4. Many research works with natural soils have shown that SSA is strongly related to the amount and kind of clay. SSA can be considered as a characteristic intrinsic soil property. It exerts strong influence on the engineering behavior of fine-grained soils (Lutenegger and Cerato, 2001; Cerato and Lutenegger, 2005). A linear relation of clay fraction and SSA of the soils from Vietnam is plotted

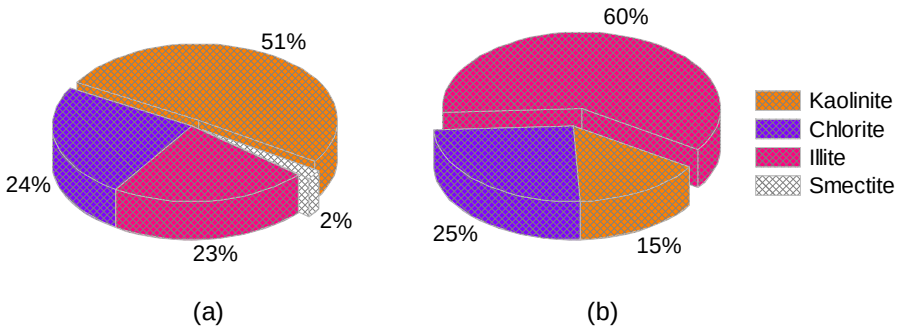


Figure 3.16: Clay mineral contents of typical clayed soils from Germany and Vietnam: (a) Specimen XB483 from Loebnitz, Germany, with dominant kaolinite accounting for 51%; (b) Specimen XB511 from Tra Linh, Thai Binh, Vietnam, with dominant illite accounting for 60%.

in comparison with the data of other authors in Figure 3.17. The resulting fitting coefficients are tabulated in Table 3.5. The relation of physical and chemical properties of various type of soils in Israel was investigated by Banin and Amiel (1970). Montomollionite was identified to be the prevailing clay fraction. A high content of Calcite ($CaCO_3$) was found in those soil samples. The presence of dominant montmollionite in the soils resulted in high SSA. This is a reason why the soils from Israel result in the steepest slope as shown in Figure 3.17. Warkentin (1968), Warkentin (1972), and Mortland (1974) investigated the clay content and SSA relations of various soils in the U.S.A. and Canada with certain amount of montmollionite. The resulting linear relations run nearly in parallel as shown in Figure 3.17. It can be figured out that the predominant kind of clay mineral in soils of these locations is identical. The soils from the Red River Delta, Vietnam, and those from Northern Adriatic and Venice, Italy (Rabitti et al., 1983) are characterized by similar depositional conditions of alluvial marine environment (coastal zones). Illite mineral is identified to be the

Table 3.3: Semi-quantitative clay mineralogical analysis of soils.

No.	Sample	Mineral content				Location
		Illite	Kaolinite	Chlorite	Smectite	
1	XB489	+++	+	+	-	DL, VN
2	XB492	+++	++	+	-	DL, VN
3	XB493	++	++	++	-	DL, VN
4	XB498	+++	++	+	-	VD, VN
5	XB511	+++	+	+	-	TL, VN
6	XB523	+++	+	++	-	YD, VN
7	XB524	+++	+	+	-	YD, VN
8	XB923	+++	+	++	-	AB, VN
9	XB450	++	+++	++	-	LN, G
10	XB459	++	+++	++	-	LN, G
11	XB460	++	+++	++	-	LN, G
12	XB476	++	++	+	-	FR, G
13	XB483	++	+++	++	+	FR, G

N.B.: +++ = high content; ++ = medium content; + = Low content;
Site Locations: DL: Dong Lam, Vietnam; VD: Vu Doai, Vietnam;
TL: Tra Linh, Vietnam; YD: Yen Dinh, Vietnam; AB: An Bai, Vietnam;
LN: Löbnitz, Germany; FR: Friedersdorf, Germany.

Table 3.4: Specific surface area (SSA) and associated clay fraction of soils.

No.	Sample	Weight	Clay fraction	SSA	Location
		(g)	(%)	(m^2/g)	
1	XB489	7.52	7.00	2.58	DL, VN
2	XB492	7.95	16.00	5.33	DL, VN
3	XB493	8.33	24.00	8.04	DL, VN
4	XB498	9.59	22.00	10.15	VD, VN
5	XB907	7.67	27.00	12.53	AB, VN
6	XB923	9.55	23.00	6.64	AB, VN
7	XB927	6.20	25.00	15.18	AB, VN
8	XB930	6.66	32.00	17.81	AB, VN
9	XB948	7.28	14.00	7.00	AB, VN
10	XB949	9.77	33.00	20.47	AB, VN

N.B.: Site locations: DL, VN: Dong Lam, Vietnam;
VD, VN: Vu Doai, Vietnam; AB, VN: An Bai, Vietnam.

predominant clay mineral. Thus, it becomes obvious why the SSA and clay fraction relations of the two sites are in parallel and closed to each other.

Farrar and Coleman (1967) determined the relation between SSA and liquid limit (LL) of British clay soils given as LL= 19 + 0.56 SSA with SSA in m^2/g and LL in %. Dolinar and Trauner (2004) found the relation as LL = 31.9 + 0.81 SSA. The liquid limit of a soil primarily depends on the type and quantity of clay minerals which are well indicated by SSA. In this study, the fine-grained soils with the same type of dominant clay mineral, illite, the dependence of liquid limit on specific surface area

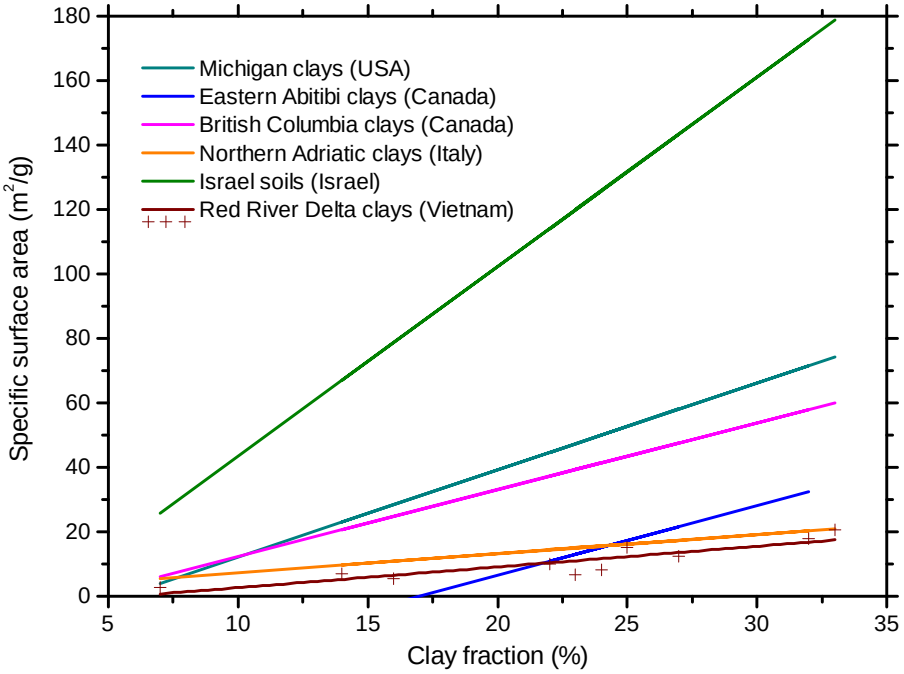


Figure 3.17: Specific surface area versus clay fraction.

Table 3.5: Linear regression of clay fraction and specific surface area with α being the slope and β the intercept.

Origin of soil samples	Coefficients		References
	α	β	
Michigan clays, USA	2.70	-14.86	Mortland (1974)
Eastern Abitibi clays, Canada	6.67	16.67	Morin and Jacobs (1984)
British Columbia clays, Canada	2.08	-8.54	Warkentin (1972)
Northern Adriatic clays, Italy	0.59	1.44	Rabitti et al. (1982)
Israel clays, Israel	5.88	-15.29	Banin and Amiel (1970)
Red River Delta clays, Vietnam	0.64	-3.77	New data

is determined as LL = 30.16 + 0.74 SSA.

The large specific surface area of clays presents more contact area between particles

Table 3.6: Linear regression of specific surface area and liquid limit

Origin of soil samples	Coefficients		References
	α	β	
British clays, Great Britain	0.56	19.00	Farrar and Coleman (1967)
Alberta clays, Canada	0.37	18.00	Warkentin (1968)
Natural clays, India	0.31	24.25	Sridharan and Prakash (1998)
Israel clays, Israel	0.14	19.81	Smith et al. (1985)
Pure clay particles	0.81	31.90	Dolinar and Trauner (2004)
Red River Delta clays, Vietnam	0.77	29.23	New data

as well as more opportunity for various inter-particle forces to develop. It also

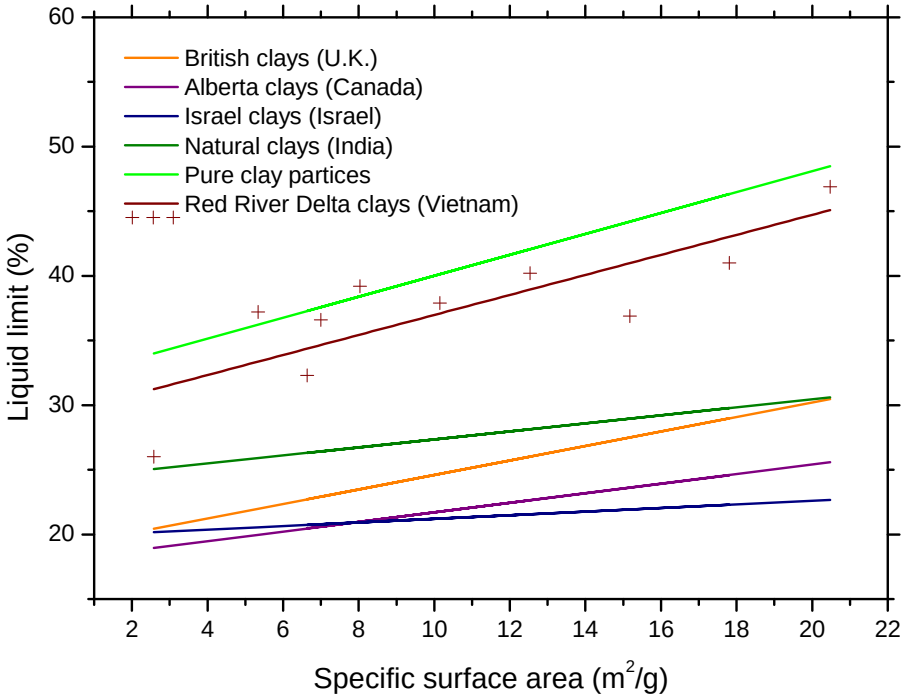


Figure 3.18: Relation between specific surface area and liquid limit with α being the slope and β the intercept.

provides more area for water molecules to be attached, thus giving clays a much greater affinity for absorbing water. The interactions between this water and the clay minerals are quite complex and beyond this study, but the ultimate effect is that the engineering properties of soils alter as the water content varies.

3.6 Petrophysical properties of soils

The petrophysical properties of soils from various site locations of Vietnam and Germany were extensively investigated in laboratory. The soil characterization of spectral induced polarization (SIP), dielectric permittivity, magnetic susceptibility and grain density were performed on more than 100 of soil specimens. Figure 3.19 illustrates the behavior of electrical conductivity of typical clay and sand samples from Germany and Vietnam in the low frequency range of 10^{-3} to 10^{3} Hz. The upper

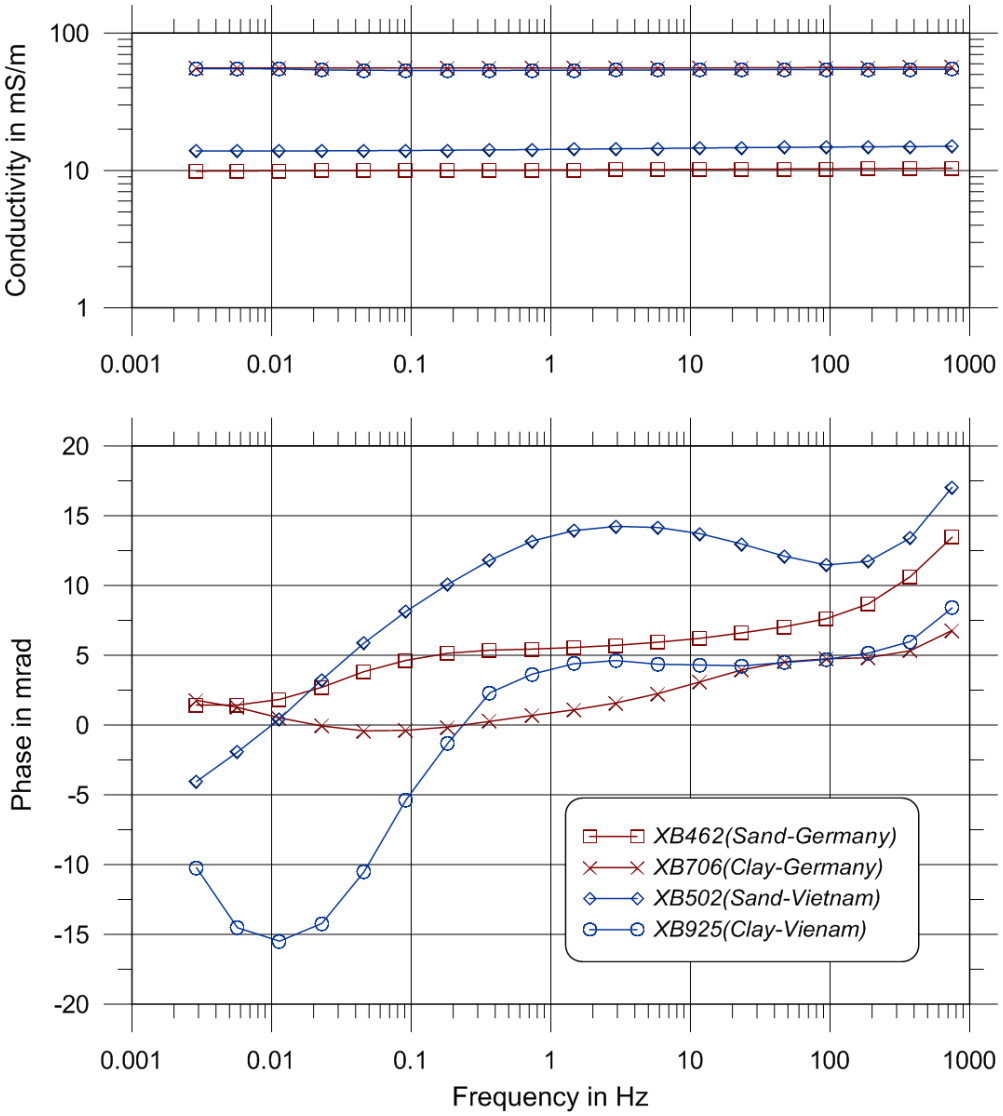
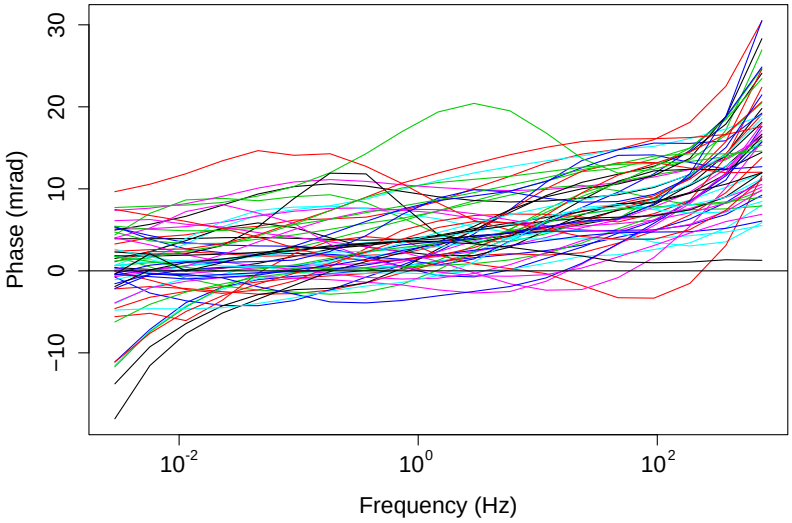


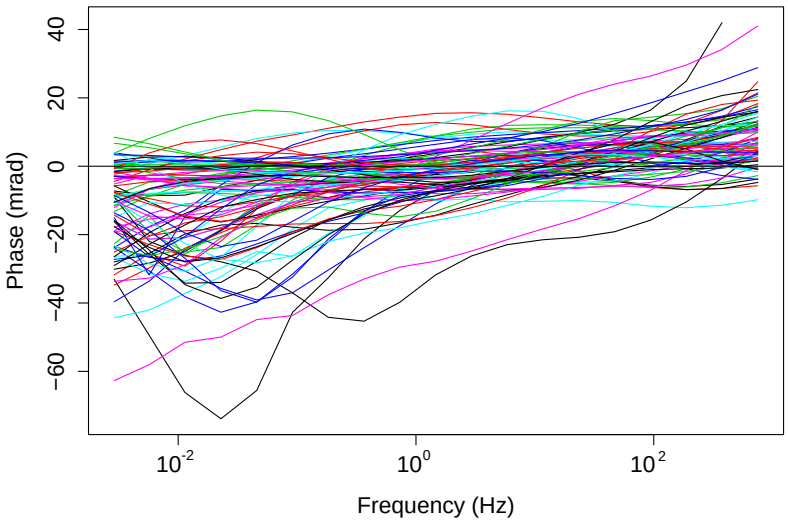
Figure 3.19: The complex conductivity spectra of typical clay and sand samples from Vietnam and Germany.

panel presents the conductivity amplitude of sand and clay where the values are different and keep constant among frequencies. The lower panel shows the behavior of the phase shift which approximately corresponds to the ratio between imaginary and real components of conductivity, among frequencies. Within the lowest frequency range, less than 1 Hz, the negative phase values usually occur in soil specimens from Vietnam. The complex conductivity of a soil when subjected to an applied current results from the movement of ions through the soil medium. The ions may

either propagate through the ionic solution (pore water) in the soil pores (electrolytic conduction) or along surfaces of clay minerals in the soils (interface conduction). The interface conduction in combination with electrochemical process involving cation exchange capacity (CEC) of the soil's clay minerals and pore water lead to a frequency dependent conductivity or spectral induced polarization.



(a) Phase spectra of soils from Germany



(b) Phase spectra of soils from Vietnam

Figure 3.20: Spectral phase shift behaviors of soils over low frequency range

Brandes (2005) reported that the negative chargeability or negative phase shift of soil results from a nonlinear behavior of current and voltage when high clay content is present. The effects of clay content and pore electrolyte concentration on the the spectral electrical response of clay-sand mixtures and unconsolidated earth materials were observed by Boadu and Seabrook (2006), Boadu and Owusu-Nimo (2010) and others. The changes of clay content in the soils result in characteristic changes in the amplitude and phase spectra of the electrical response but no negative phases have been observed.

Many natural clayey soil samples from Vietnam exhibit negative IP response in a low frequency range (smaller than 1 Hz). Some soils from Germany also present negative phase but smaller in scale as shown in Figure 3.20. Some soil specimens from Vietnam show extreme negative phase values. The minimum peak approaches -75 mrad. Olhoeft (1985) has observed a similar phenomenon with a sample from an oil well with negative IP peak of -96.3 mrad. Jones (1997) investigated the interaction between pure clay minerals of montmorilionite, kaolinite with organic matters, which are available in waste sites, such as ethylene glycol, phenol, tetrachloroethylene (PCE) and trichloroethylene (TCE) by nonlinear complex resistivity analysis. However, no significant negative phase had been observed.

Two soil samples from Vietnam with extreme negative phase values were sent to organic chemical laboratory to analyse chemical constituents. In an attempt of explaining this phenomenon, several techniques are used to determine and separate the chemical components of the soils such as nuclear magnetic resonance spectroscopy (NMR), infrared spectroscopy (IR), gas chromatography-mass spectroscopy (GS-MS). The results indicate that the CH_2-groups, alkyl chains or unsaturated hydrocarbon are present in the soils. Further information couldn't be gained because the concentration is too low. The extreme negative IP effect would result from the reaction between clay minerals in soils and organic matter.

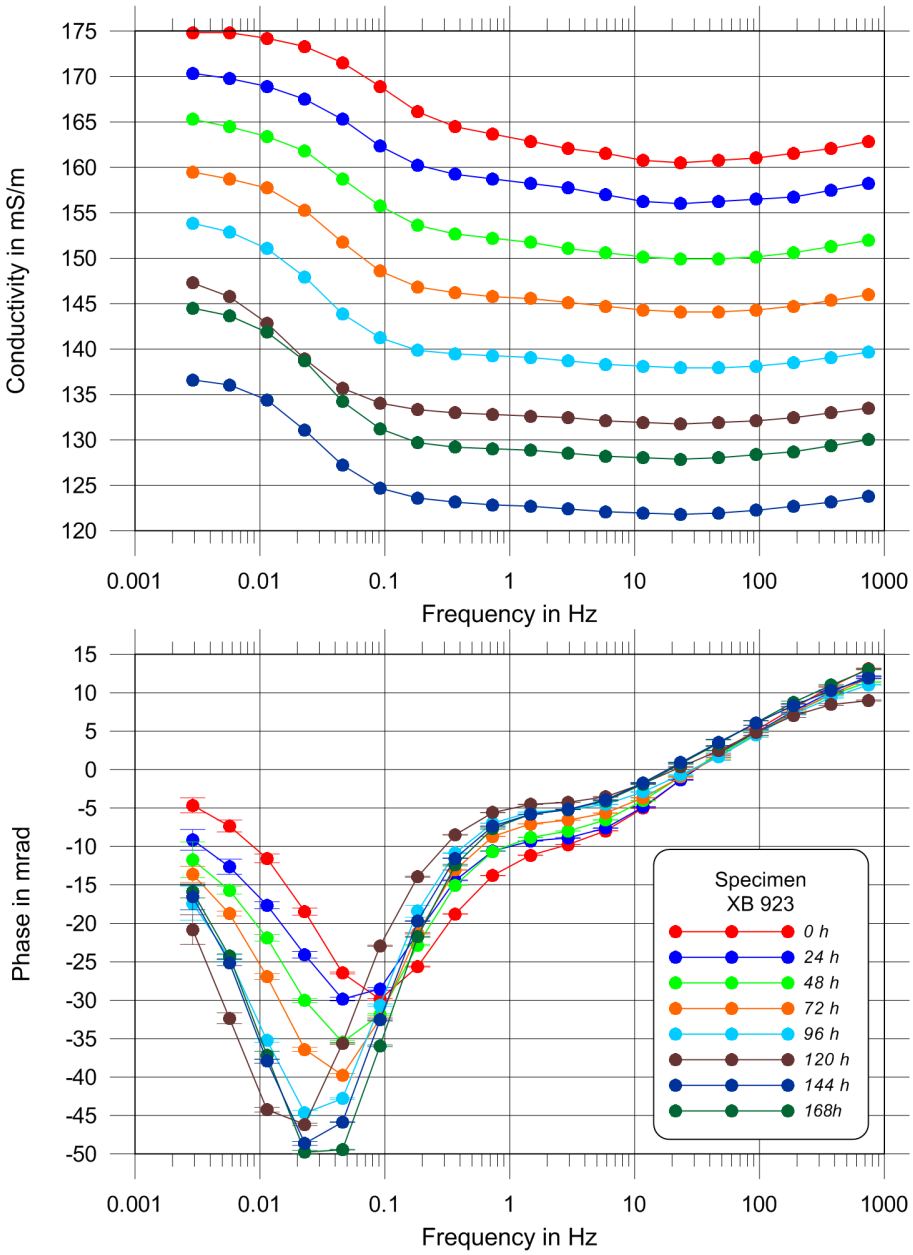


Figure 3.21: The complex resistivity spectrum of a silt soil simulating clay-organic reaction. Soil sample XB923 at the depth of 17 m, An Bai, Thai Binh, Vietnam.

Figure 3.21 illustrates clay and organic reaction during time. The measurement were repeatedly measured after every 24 h. The negative phase peak approaches initially -30 mrad at the frequency 0.1 Hz. After 148 h, the phase shift seems to

appear stable and shows a negative value of -50 mrad at 0.02 Hz and keeps almost unchanged with frequency greater than 10 Hz.

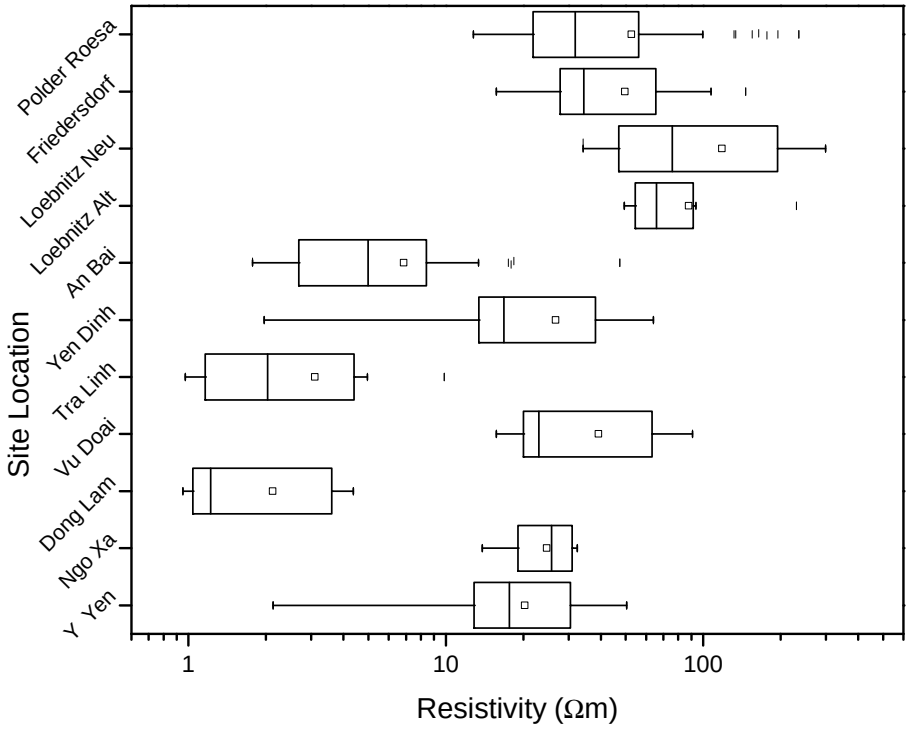


Figure 3.22: Distribution of soil resistivity at 1.4 Hz from different locations in Germany and Vietnam.

Figure 3.22 statistically summarizes the resistivity of soils at 1.4 Hz from various locations in Germany and Vietnam. The resistivity values out of range 1.5 interquartile range are considered as outlieners and symboled as a small vertical bar.

In general, soils from Germany show higher resistivity than those in Vietnam. Higher clay content and the presence of organics in the soils from Vietnam may account for lower resistivity. Especially the locations An Bai, Tra Linh and Dong Lam have very low resistivity from 1 to 6 Ωm in individual average.

Figure 3.23 shows the plot of clay fraction and resistivity with reference to water content. The general trend confirms that a higher clay fraction results in a lower resistivity except some soil samples with high salt concentration originating from

coastal areas such as the sandy soils XB 489, XB 490, XB 495 from Tien Hai, Thai Binh, which have low resistivity of 4 to 5 Ωm. In this area, some clayey soils XB 492, XB 493, XB 494 have very low resistivity of about 1 Ωm. The other outliers of the trend are stiff clays in purple color. As mentioned in Figure 3.9, these soils are located in a greater depth of 25 m belonging to an impervious aquitard layer (Do, 1996).

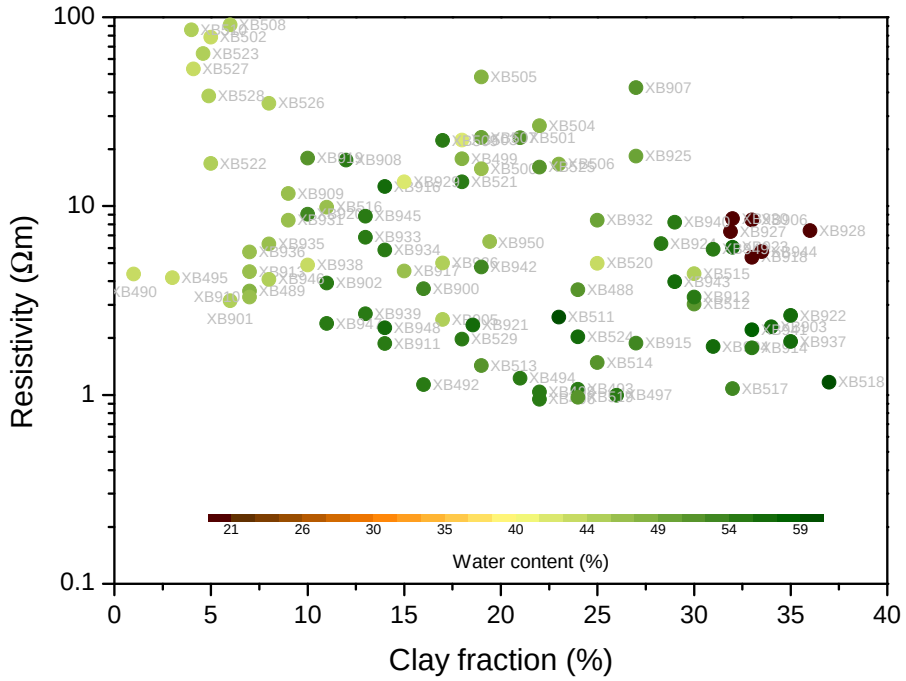


Figure 3.23: Resistivity versus clay fraction with a reference to water content.

The dielectric permittivity of soils depends upon various parameters such as frequency, water content, density, electrical conductivity, clay mineralogy and clay content (Salat and Junge, 2010). Water content and frequency have the strongest influence on soil permittivity. Figure 3.24 presents the distribution of real part of relative dielectric permittivity of soils at 200 MHz from various locations in Germany and Vietnam. The peat soils from Uchter Moor, Germany with very high water content result in very high relative permittivity. Two locations An Bai and Dong Lam are characterized by higher values compared to the other sites.

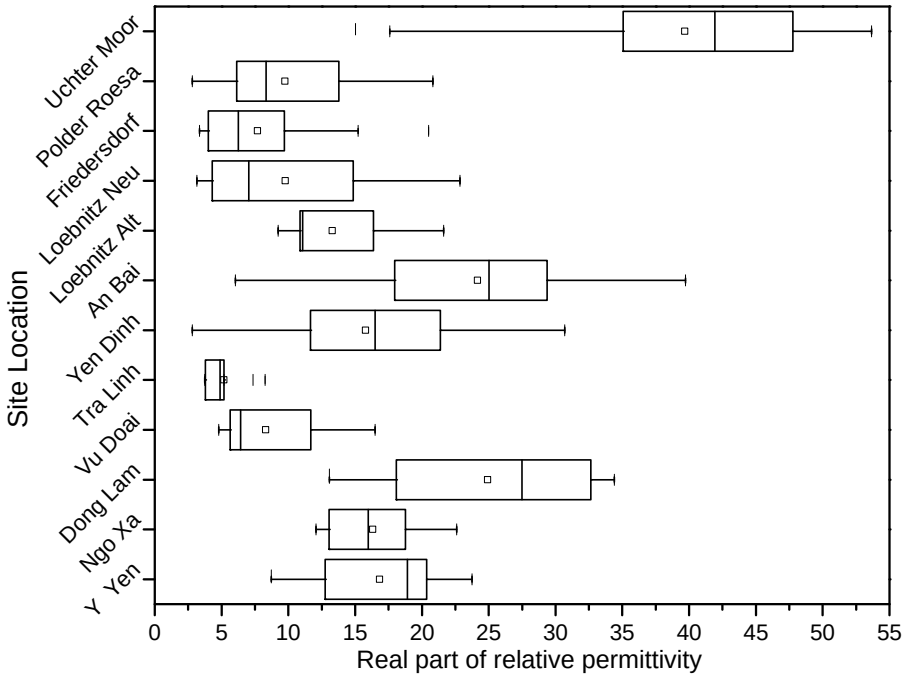


Figure 3.24: Distribution of the real part of relative permittivity, $\epsilon_r^{'}$, of soils at frequency 200 MHz.

Volumetric water content has strong effect on real part of permittivity as found in the literature. Topp et al. (1980) investigated four mixtures of soil by TDR (Time Domain Reflectometry) method and has established an empirical formula relating the volumetric water content to the apparent dielectric permittivity. Salat and Junge (2010) used parallel plate cell method to measure dielectric permittivity of soils from Rio Palancia watershed, Spain. Figure 3.25 show the relations between volumetric water content and real part of relative permittivity at 200 MHz of soils from Germany, Vietnam and others as cubic polynomial forms

$$\epsilon_r^{'} = \alpha + \beta w^* + \gamma w^{*2} + \zeta w^{*3}. \tag{3.8}$$

The three coefficients of the cubic equation are compiled in Table 3.7. The Salat and Junge's curve appears slightly higher compared than the others. Salat and

Junge (2010) explained that the soils contain a high content of carbonate, that might increase the measured dielectric permittivity.

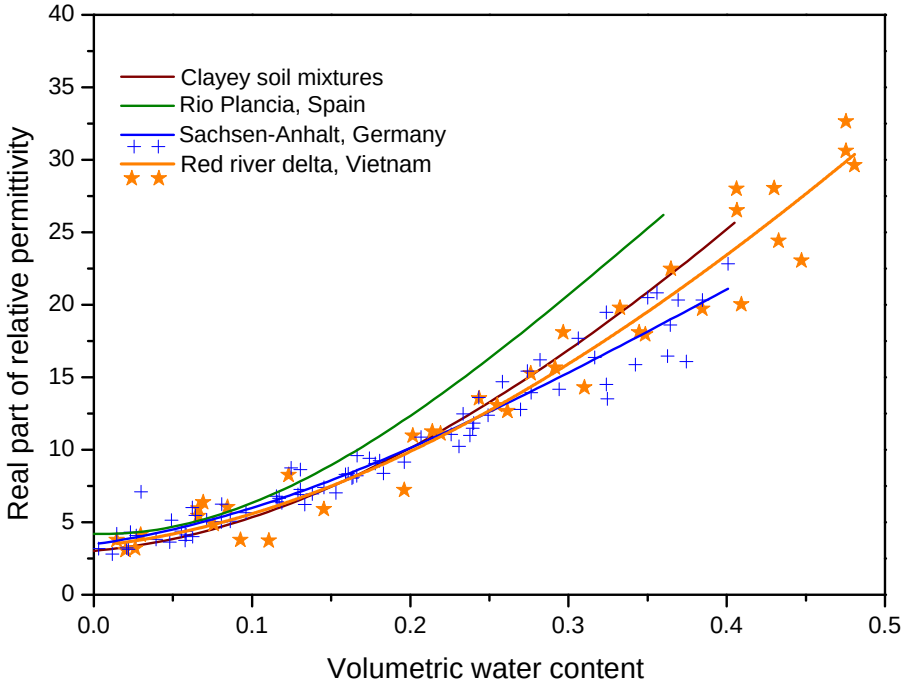


Figure 3.25: Relation of volumetric water content, w^*, and real part of relative permittivity, $\epsilon_r^{'}$, of soils at frequency 200 MHz.

Table 3.7: Cubic fitting parameters of volumetric water content and real part of relative permittivity of the equation 3.8.

Original soil samples	Coefficients				References
	α	β	γ	ζ	
Soil mixtures	3.03	9.30	146.00	-76.70	Topp et al. (1980)
Rio Palancia, Spain	4.20	-3.10	269.70	-253.70	Salat and Junge (2010)
Sachsen-Anhalt, Germany	3.48	15.31	107.39	-89.53	New data
Red River Delta, Vietnam	3.97	-3.57	202.30	-181.90	New data

The imaginary part of relative permittivity of soils from two sites of An Bai and Dong Lam is also higher compared to the others as indicated in Figure 3.26. Imaginary part of relative permittivity is proportional to the real part of conductivity (see chapter 2) and this high imaginary part of relative permittivity can be explained by very low resistivity of soils from these two sites as shown in Figure 3.22. As mentioned above, the soils of these sites may have higher clay content and organic

matter present. The imaginary part of relative permittivity of soils from Vietnam is also slightly higher compared to the soils from Germany.

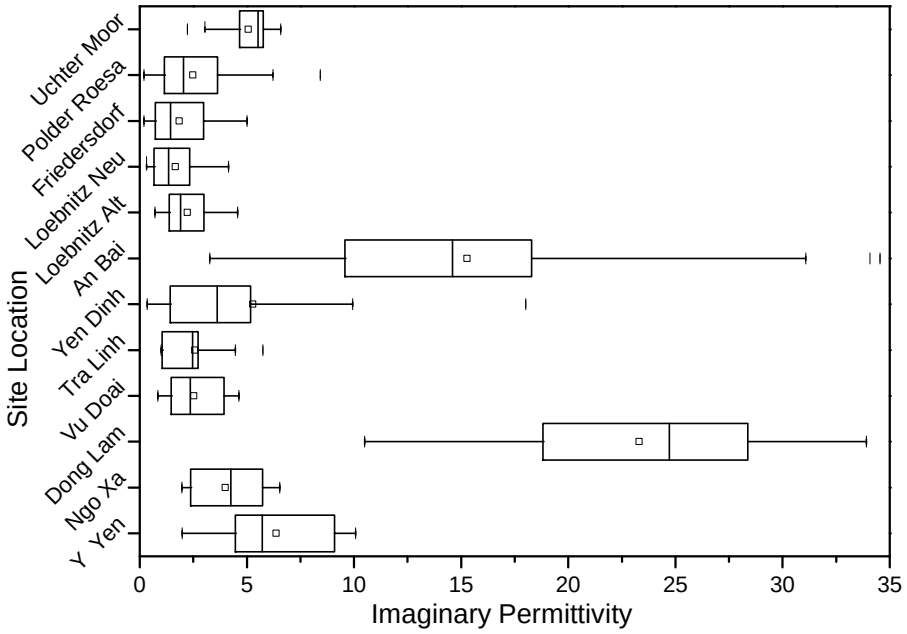


Figure 3.26: Distribution of imaginary permittivity of soil at frequency 200 MHz

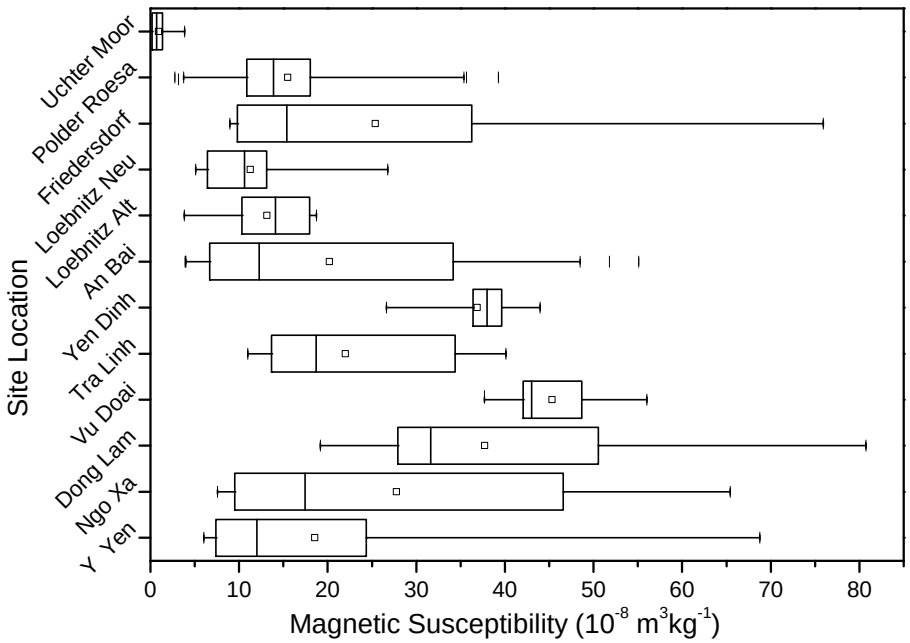


Figure 3.27: Distribution of magnetic susceptibility of soil from Germany and Vietnam

The distribution of magnetic susceptibility of soils from Germany and Vietnam is presented in Figure 3.27. The magnetic susceptibility of soils from some locations in Vietnam exhibits higher than those in Germany. From clay mineralogical analysis, the soils are mainly composed of quartz, illite, kaolinite chlorite. The typical soils from Vietnam have illite as a dominant clay mineral, while kaolinite is predominant in the soils from Germany as illustrated in Figure 3.16. Kaolinite classified as diamagnetic mineral has magnetic susceptibility of - 1.9 SI unit, where as illite is defined as paramagnetic mineral with magnetic susceptibility of 15 SI unit as shown in Table 2.2. This may be the reason accounting for a higher magnetic susceptibility in the soils from Vietnam. An other reason may be explained by the presence of more iron oxide as indicated in the clay mineralogical analysis.

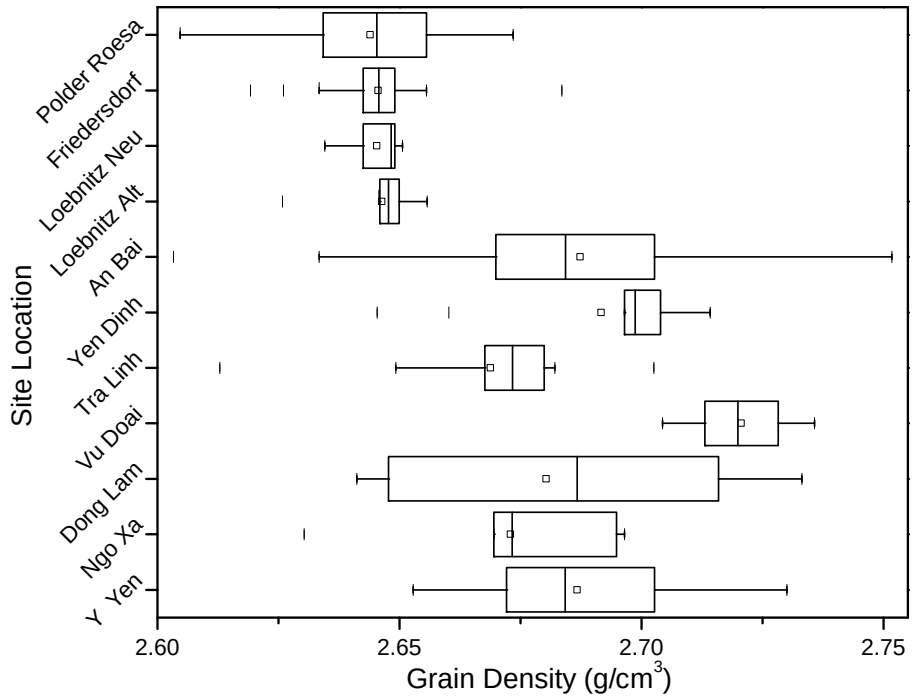


Figure 3.28: Distribution of grain density of soil from Germany and Vietnam

Two methods are widely used to determine experimentally specific gravity. One is the liquid submersion technique and the other is the gas pycnometer technique.

The gas pycnometer test method is described by ASTM D5550 on specific gravity of soil solids by gas pycnometer. Grain density and specific gravity in soil mechanics are interchangeably used. Typical values of specific gravity of general soil vary from 2.65 to 2.85 (Lambe and Whitman, 1969), while average value for general clay 2.72, for sands 2.67, for quartz 2.65, for illite 2.6 - 3.0 and for kaolinite 2.6 - 2.68 (Mitchell and Soga, 2005). As visible in Figure 3.28, the soils from German sites which kaolinite as the predominant clay mineral show an average grain density of 2.65 g/cm^3. The soils from Vietnam with the dominant clay mineral illite show an average grain density from 2.67 to 2.72 g/cm^3.

Chapter 4

Multivariate statistical analysis

4.1 Introduction

Petrophysical and geotechnical parameters are characterized by a complicated interrelationships. Most parameters depend on a variety of influencing factors. Multivariate statistical analysis is a promising tool to investigate the interrelationship between different parameters. It becomes necessary to analyze comprehensive data sets of soil samples in simultaneous manner to figure out relations between the parameters in question. In geosciences, some authors have well developed and presented the methods of classical multivariate statistics such as Swan and Sandilands (1995), Reyment and Savazzi (1999), Davis (2002), Wackernagel (2003). The requirement for a fresh approach of multivariate statistics is caused by three recent development: (i) many classical methods exhibit poor results when dealing with large and complex data sets; (ii) the questions on a large data set is different from those of a small data set in previous time; and (iii) numerous recent improvements in computational power and equipments.

The aim of this chapter is to present the state of the art at methods that have been developed in an attempt to discover any hidden patterns or structures of the data set. The goal of experimental science is the understanding and exploration of

unknown relations including natural laws. To explore these laws, the methods of classification, comparison and conjecture or prediction are used. In classification, relevant parameters have to be measured. Methods of hypothesis tests and statistical models will be used for comparison and conjecture. We are familiar with a display of scatterplot of two variables. However, when dealing with more than two variables, it becomes more difficult to visualize the multivariate data distributions.

4.2 Data structure and proximity measures

4.2.1 Data structure

A starting point for all multivariate statistical algorithms is a matrix in which rows represent the objects and columns the variables. In this study, an object represents an investigated soil sample. As a matter of convention, the rows of a data matrix present soil samples and the columns present the properties of this sample. In multivariate statistics the number of samples n should be larger than the number of properties p. The data matrix is of rectangular shape $n \times p$. The various properties of a sample are normally distributed in the range of magnitude. The manipulation of centering or normalizing of the data matrix should be done prior to the multivariate statistical analysis.

4.2.2 Measurement of proximity

Measures of proximity are of two types: similarity and dissimilarity (distance) with the obvious interpretation of a measure that indicates how similar or dissimilar objects are to each other. Upon the type of data, several authors discuss on various similarity and dissimilarity definition together with associated problems (Gower, 1985, Baulieu, 1989, Jackson et al., 1989 and Gordon, 1999). The choice of a proximity

measure depends upon the problem at hand. More details can be found in Sneath and Soka (1973). Only some widely used distances in geosciences are briefly given here.

Suppose a data set of n objects has dissimilarities δ_{rs} measured between all pairs of objects. A configuration of n points representing the objects is sought in a p dimensional space. Each point represents one object, with the r^{th} point representing object r. Let the distances between pairs of points be d_{rs}. The aim of visualization is to find a configuration such that the distances d_{rs} match as well as possible the dissimilarities δ_{rs}. The different notions of "matching" give rise to different techniques of multidimensional scaling (MDS).

Euclidean metrics

These metric measures correspond to the straight line distances in Euclidean space. In a univariate example, the Euclidean distance between two values is the arithmetic difference. In a bivariate case, the minimum distance between two points is the hypotenuse of the right-angled triangle in the two-dimensional space. For three variables the hypotenuse extends through three-dimensional space. Although it is difficult to visualize, an extension of Pythagoras theorem gives the Euclidean distance between two points in n-dimensional space:

$$\delta_{rs} = \left\{ \sum_{i=1}^{p} (x_{ri} - x_{si})^2 \right\}^{1/2} . \tag{4.1}$$

A weighted Euclidean distance considers varying weights for different properties:

$$\delta_{rs} = \left\{ \sum_{i=1}^{p} w_i (x_{ri} - x_{si})^2 \right\}^{1/2} . \tag{4.2}$$

Mahalanobis distance

The Mahalanobis distance is a generalised form of an Euclidean distance which

weights variables using the sample variance-covariance matrix $\mathbf{R}$. Because the covariance matrix is used this also means that correlations between variables are taken into account. The Mahalanobis distance is normally used to measure the difference between the means of two multivariate groups. It can be defined as a similarity measure between two vectors $\mathbf{x}_r$ and $\mathbf{x}_s$ of the data matrix with the covariance matrix $\mathbf{R}$:

$$\delta_{rs} = \left\{[\mathbf{x}_r - \mathbf{x}_s]^T \mathbf{R}^{-1} [\mathbf{x}_r - \mathbf{x}_s]\right\}^{1/2}. \tag{4.3}$$

Cosine coefficient

The cosine coefficient expresses the dissimilarity between object a and object b by regarding each as a vector defined in a p dimensional space. This is a non-Euclidean, pattern similarity metric. The cosine of the angle between two vectors is identical to their correlation coefficient. However, unlike a normal correlation calculation the pairs of values are drawn from different variables for two cases rather than two variables from different cases. Cosine of two object a and b is defined as

$$cosine\ \theta_{ab} = \frac{\sum_{k=1}^{p} x_{ak}\ x_{bk}}{\left\{(\sum_{k=1}^{p} x_{ak}^2)(\sum_{k=1}^{p} x_{bk}^2)\right\}^{1/2}}. \tag{4.4}$$

4.3 Cluster Analysis

Cluster analysis and classification both are techniques of placing objects into groups or classes. The difference is that in a cluster analysis the classes are not predefined as in classification. Cluster analysis, which is the most well-known example of unsupervised learning, is a very popular tool for analyzing unstructured multivariate data. The methodology consists of various algorithms each of which seeks to organize a given data set into homogeneous subgroups, or "clusters". A cluster is simply a

collection of samples that are more "similar" to each other than they are to samples in other clusters. There is no guarantee that more than one such group can be found. In any practical application, however, the underlying hypothesis is that the data form a heterogeneous set should be separated into natural groups.

All clustering algorithms begin by measuring the similarity or dissimilarity between the samples to be clustered. Similar samples will be placed into the same cluster. It is also possible to view similarity by its inverse, the distance between cases, with distance declining as similarity increases. This leads to a general conclusion that objects in the same cluster will be closer to each other or more similar than they are to objects in other clusters. It also means that there must some means of measuring distance. The most obvious distances are Euclidean which are straight lines that can be measured with a "ruler" while others, often based on similarity, are non-Euclidean.

4.3.1 Partitioning methods

Partitioning techniques encounter the problem of dividing n samples, described by p variables, into a small number k of discrete classes. The k-means is one of the simplest unsupervised learning algorithms that solve the known clustering problem. The most intuitive and frequently used criterion function in partitioned clustering techniques is the squared error criterion or objective function. This algorithm aims at minimizing an objective function

$$J = \sum_{j=1}^{k} \sum_{i=1}^{n} \left\| x_i^{(j)} - c_j \right\|^2, \tag{4.5}$$

where $\left\| x_i^{(j)} - c_j \right\|^2$ is a chosen distance measure between a data point $x_i^{(j)}$ and the cluster center c_j. The objective function J is an indicator of the distance of the n data points from their respective cluster centers. The algorithm is composed of the following steps that are iterated until a solution is found.

1. Choose k cluster centers that coincide with k randomly defined points. The initial clusters could be random or based on some "seed" values.

2. Repartition by assigning each samples to the closest cluster center.

3. Recompute the cluster centers as centroids.

4. Repeat steps two and three until convergence is achieved. The endpoint will be the minimum of the objective function J.

The k-means algorithm does not necessarily find the optimal configuration, corresponding to the global objective function minimum. The algorithm is also significantly sensitive to the initial randomly selected cluster centers. The k-means algorithm can be run multiple times to reduce this effect.

4.3.2 Hierarchical method

Hierarchical techniques are the most widely applied clustering techniques in the earth sciences. Gordon (1987) reported an excellent review on hierarchical classification. This method joins the most similar observations, then successively connects the next most similar ones to these. The graphic display of the complete clustering process is a dendrogram. The nodes of the dendrogram represent clusters, and the lengths of the stems (heights) represent the distances at which clusters are joined.
Given a set of n items to be clustered, and an $n \times n$ distance matrix, the basic process of hierarchical clustering runs as follows:

1. Assign each item to a cluster, each cluster contains just one item. Let the distances between the clusters be the same as the distances between the items they contain.

2. Find the closest or most similar pair of clusters and merge them into a single cluster, so that now we have one cluster less.

3. Calculate the distances between the new cluster and each of the old clusters.

4. Iterate steps 2 and 3 until all items are clustered into a single cluster of size n.

Step 3 can be done in different ways, which is what distinguishes single-linkage from complete-linkage and average-linkage clustering. In single-linkage clustering, we consider the distance between one cluster and another cluster to be equal to the shortest distance from any member of one cluster to any member of the other cluster. If the data consists of similarities, we consider the similarity between one cluster and another cluster to be equal to the greatest similarity from any member of one cluster to any member of the other cluster. In complete-linkage clustering, we consider the distance between one cluster and another cluster to be equal to the greatest distance from any member of one cluster to any member of the other cluster. In average-linkage clustering, we consider the distance between one cluster and another cluster to be equal to the average distance from any member of one cluster to any member of the other cluster.

This type of hierarchical clustering is called agglomerative because it merges clusters iteratively. There is also a divisive hierarchical clustering which does the reverse by starting with all objects in one cluster and subdividing them into smaller pieces. Divisive methods are not generally available, and have been rarely applied in earth sciences.

4.4 Principal Component Analysis (PCA)

Principal Component Analysis (PCA) is the most widely used method of multivariate data analysis due to the its simple algorithm and straightforward interpretation. The major goal of PCA is to reveal hidden structures in a data set. In geosciences, PCA can be used for (i) Reducing the dimensionality of the data, (ii) Multivariate outliers

detection, (iii) Decoding a correlation matrix, (iv) Identifying underlying factors, (v) Detecting intrinsic correlation, and (vi) Preparing the data for further analysis using other techniques (Jolliffe, 2002, Wackernagel, 2003).

The algebraic solution to PCA can be easily found in any multivariate analysis textbooks such as Jolliffe (2002), Davis (2002), Wackernagel (2003), or Hair and Anderson (2010). Therefore, only a brief description of PCA is mentioned here.

Consider a data matrix $\mathbf{X}_{n\times p}$ of n soil samples and p variables. If the variances are significantly different in the data or variables measured in different dimensions, the data matrix should be standardized by subtracting the means of each row and scaling each row by dividing by its standard deviation. The variance-covariance matrix can be calculated as

$$\mathbf{R} = \mathbf{X}^T\mathbf{X}. \tag{4.6}$$

The interrelationships between a data matrix and the eigenvalues and eigenvectors of its two cross product matrices are expressed in the singular value decomposition (SVD), well known as Eckart-Young theorem:

$$\mathbf{X} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T, \tag{4.7}$$

with $\mathbf{U}_{n\times n}$ and $\mathbf{V}^T_{p\times p}$, the transpose of $\mathbf{V}$, being unitary matrices. $\mathbf{\Sigma}_{n\times p}$ is a diagonal matrix with non-negative numbers on the diagonal containing the singular values of $\mathbf{X}$. The columns of $\mathbf{V}$ are termed the principal components or the principal component loadings in PCA literature.

The fundamental characteristic of PCA is to approximate $\mathbf{X}$ by a lower rank matrix $\hat{\mathbf{X}}$ which minimizes the residual distance $\|\mathbf{X}-\hat{\mathbf{X}}\|$ on the basis of an approximation of the least squares criterion. The r dimensional Eckart Young approximation becomes more informative as

$$\hat{\mathbf{X}} = \mathbf{U}\mathbf{\Sigma}_r\mathbf{V}_r^T = \mathbf{X}\mathbf{V}_r\mathbf{V}_r^T, \tag{4.8}$$

where r, the lower rank. The rows of $\mathbf{U}\mathbf{\Sigma}_r$ give the r dimensional coordinates for the n samples, while the columns of $\mathbf{V}^T$, the rows of $\mathbf{V}$, give the directions of the biplots axes which will be elaborated in the next sections.

To approximate a data matrix $\mathbf{X}$, the analysis is based on the singular value decomposition (SVD), while an approximation of the variance-covariance matrix $\mathbf{X}^T\mathbf{X}$ is based on the spectral eigenvalues decomposition which happens to coincide with its singular value decomposition. The variance-covariance matrix can be written as

$$\mathbf{R} = \mathbf{X}^T\mathbf{X} = \mathbf{V}\mathbf{\Sigma}^2\mathbf{V}^T, \tag{4.9}$$

and the approximation to the variance-covariance matrix is given as

$$\hat{\mathbf{X}}^T\hat{\mathbf{X}} = \mathbf{V}_r\mathbf{\Sigma}_r^2\mathbf{V}_r^T. \tag{4.10}$$

The matrix $\mathbf{V}_r\mathbf{V}_r^T$ in equation 4.8 is considered as the projection matrix. The rows of $\mathbf{X}$ and $\mathbf{V}_r\mathbf{V}_r^T$ result in the projections of $\mathbf{X}$ onto the r dimensional subspace relative to the original p orthogonal axes.

The PCA method is actually a statistical interpretation of the eigenvalues. Multiplying the data matrix $\mathbf{X}$ with the eigenvector matrix $\mathbf{V}$ results in a score matrix $\mathbf{Y}$ that contains the principal components:

$$\mathbf{Y} = \mathbf{X}\mathbf{V} = \mathbf{U}\mathbf{\Sigma}. \tag{4.11}$$

4.4.1 PCA visualization in multidimensional space

For the graphical representation, a PCA approximation is presented in a r-dimensional subspace $\mathfrak{L}$ of the p space which results from a best fitting in the least squares sense. With the r-dimensional subspace $\mathfrak{L}$, n points can be orthogonally projected on it.

This subspace is characterized by the minimum sum of squares residuals between the original points and their projections. When representing the samples relative to orthogonal axes in $\mathfrak{L}$, the coordinates of the projected points are given by

$$\mathbf{Y} = \mathbf{X_r}\mathbf{V_r}. \tag{4.12}$$

If the dimension $r = 2$, the best fitting subspace $\mathfrak{L}$ will be a plane of best fit. The projection of original axes on the r-dimensional best fitting subspace defines the biplot axes.

4.4.2 Measures of fit in PCA method

The overall quality of approximation of the sample matrix $\mathbf{X}$ is usually measured as the ratio of the variance on the corresponding approximation to the total variance or in terms of fitted to total sums of squares:

$$\frac{Variance\ of\ the\ factor}{Total\ variance} = \frac{\Sigma_{i=1}^{r}\sigma_i^2}{\Sigma_{i=1}^{p}\sigma_i^2}, \tag{4.13}$$

where σ_i are the eigenvalues.

To measure the quality of approximating variable, a quantity termed adequacy is used to assess the approximations of the sample matrix $\mathbf{X}$. In a r-dimensional approximation, the adequacy of a specific variable is defined as

$$Adequacy = \Sigma_{i=1}^{r} v_i, \tag{4.14}$$

where v_i are the i^{th} diagonal values of $\mathbf{V}_r\mathbf{V}_r^T$. Axis adequacy is a measure of sums of squares of the rows of the eigenvectors matrix.

Axes predictivity is another important parameter to measure the fitting quality. It is defined as the ratio of the diagonal of the variance-covariance approximation matrix

to the corresponding elements of variance-covariance matrix:

$$Predictivity = diag(\hat{\mathbf{X}}^T\hat{\mathbf{X}}) \left\{diag(\mathbf{X}^T\mathbf{X})\right\}^{-1}, \tag{4.15}$$

where the terms $\mathbf{X}^T\mathbf{X}$ and $\hat{\mathbf{X}}^T\hat{\mathbf{X}}$ were described in equations 4.8 and 4.10, respectively. The details on the measures of fit in PCA analysis can be found in Gardner-Lubbe et al. (2008).

Chapter 5

Soil properties analysis using multivariate statistics

5.1 Example 1: Geotechnical data

5.1.1 Data preparation

A set of 93 soil samples described by geotechnical soil properties namely water content, density, cohesion, internal friction angle, clay fraction and porosity was analysed in simultaneous manner with principal components analysis (PCA) method. As an input preparation for the principal components analysis, the sample data set was randomly arranged in a matrix $\mathbf{X}$ of 93 rows and 6 columns (Table 5.1). The rows are the soil samples and the columns are the soil sample properties. The magnitudes of the values of soil properties are different in dimension. Differences in magnitude may distort the computation and the variables can not obviously be analysed and compared. Standardization or normalization of the sample data matrix is required prior to undertaking the PCA. The sample data matrix $\mathbf{X}_{93\times6}$ is randomly sampled and standardized by centring and scaling, by dividing each variable by its standard deviation, becoming $\tilde{\mathbf{X}}_{93\times6}$.

Table 5.1: Geotechnical properties of soils from Vietnam.

Sample	ω	Density	Cohesion	F.Angle	CF	Porosity
	(%)	(g/cm^3)	(kPa)	(deg)	(%)	(%)
XB488	34.50	1.75	8.34	7.39	24.00	51.90
XB489	32.50	1.82	3.14	18.39	7.00	48.90
XB490	26.80	1.90	3.73	23.61	1.00	43.60
XB491	27.10	1.92	2.84	23.31	2.30	43.80
XB492	42.70	1.75	6.18	6.50	16.00	54.40
XB493	39.60	1.75	6.57	6.50	24.00	54.00
XB494	41.40	1.73	6.77	6.12	21.00	54.90
XB495	27.50	1.90	1.28	24.51	3.00	44.00
XB496	40.60	1.72	7.16	6.50	22.00	54.80
XB497	41.40	1.73	6.57	6.50	26.00	55.00
XB498	39.50	1.73	7.85	5.37	22.00	54.20
XB499	30.20	1.81	12.36	13.55	18.00	48.80
XB500	28.50	1.86	14.52	14.44	19.00	46.40
XB501	37.80	1.83	6.28	8.11	21.00	51.00
XB502	27.50	1.90	2.16	20.98	5.00	44.40
XB503	21.00	1.88	9.81	16.77	18.00	42.50
XB504	33.90	1.85	3.63	13.85	22.00	49.00
XB505	31.30	1.81	11.28	10.86	19.00	49.10
XB506	30.40	1.84	8.04	10.25	23.00	47.90
XB507	33.50	1.80	2.35	6.87	19.00	50.20
XB508	28.30	1.90	2.55	19.88	6.00	44.50
XB509	42.50	1.76	8.04	5.10	17.00	54.30
XB510	29.50	1.92	0.39	21.60	4.00	44.70
XB511	58.50	1.61	4.91	4.39	23.00	62.40
XB512	37.40	1.82	12.36	9.33	30.00	51.10
XB513	37.00	1.80	7.16	7.49	19.00	51.30
XB514	36.50	1.82	8.93	10.86	25.00	50.80
XB515	29.20	1.94	10.79	17.49	30.00	44.80
XB516	29.70	1.84	3.63	15.03	11.00	47.10
XB517	40.10	1.76	5.40	6.25	32.00	53.60
XB518	52.00	1.68	5.69	7.49	37.00	59.10
XB519	36.50	1.77	6.28	9.33	24.00	52.00
XB520	25.90	1.94	12.56	15.32	25.00	43.10
XB521	41.60	1.74	5.59	6.73	18.00	54.40
XB522	28.60	1.88	2.55	19.75	5.00	45.50
XB523	28.00	1.89	1.96	20.23	4.00	44.60
XB524	44.30	1.72	5.49	5.48	24.00	55.90
XB525	37.60	1.79	6.38	7.74	22.00	51.90
XB526	29.40	1.88	2.84	19.14	8.00	45.90
XB527	27.50	1.90	1.86	20.98	4.00	44.20
XB528	28.40	1.90	2.26	19.85	4.00	44.60
XB529	42.80	1.73	5.98	5.60	18.00	55.20
XB900	39.50	1.76	5.69	7.74	16.00	53.30
XB901	29.50	1.90	2.52	19.75	6.00	45.10
XB902	43.20	1.73	4.87	6.50	11.00	55.20
XB903	46.80	1.71	8.14	6.96	37.00	56.80
XB904	44.20	1.75	8.50	6.73	31.00	55.40
XB905	27.10	1.85	14.68	15.91	19.00	44.80
XB906	24.90	1.98	18.38	16.49	33.00	41.30
XB907	35.70	1.80	8.27	9.09	27.00	50.90
XB908	42.50	1.74	4.87	5.94	12.00	54.60
XB909	30.50	1.87	2.58	20.23	9.00	46.60
XB910	31.60	1.71	3.12	19.12	7.50	46.60
XB911	43.50	1.73	5.46	6.05	14.00	55.20
XB912	40.60	1.76	8.80	6.73	30.00	53.90
XB913	30.50	1.87	3.24	19.98	7.00	46.60
XB914	42.53	1.78	7.72	7.42	31.00	55.00
XB915	34.60	1.85	10.14	9.37	27.00	49.40
XB916	45.20	1.72	5.46	6.05	14.00	56.30
XB917	28.20	1.86	16.45	13.12	15.00	46.30
XB918	24.30	1.95	20.96	16.22	33.00	42.10
XB919	38.50	1.78	5.62	8.87	10.00	52.00
XB920	39.20	1.75	4.94	6.73	10.00	53.20
XB921	44.50	1.72	5.59	5.94	17.00	55.90
XB922	46.50	1.72	8.18	6.73	35.00	56.80
XB923	44.50	1.72	6.67	6.62	32.00	56.10
XB924	41.90	1.73	7.32	7.74	29.00	55.00
XB925	35.20	1.79	9.36	9.85	26.00	49.40
XB926	27.90	1.90	14.62	15.19	17.00	44.80
XB927	23.10	2.02	18.38	16.49	25.00	39.50
XB928	23.50	1.99	20.34	16.73	36.00	40.60
XB929	25.40	1.95	14.62	14.14	15.00	42.20
XB930	24.40	2.01	17.79	16.73	32.00	40.20
XB931	24.20	1.96	8.83	18.26	9.00	41.00
XB932	34.90	1.84	8.96	9.62	25.00	49.80
XB933	40.69	1.78	5.11	6.81	15.00	54.60
XB934	40.50	1.74	5.20	6.28	14.00	53.90
XB935	31.30	1.88	3.14	19.01	8.00	46.60
XB936	30.50	1.87	3.24	19.98	7.00	46.60
XB937	46.50	1.72	8.18	6.73	35.00	56.80
XB938	25.90	1.92	9.58	15.70	10.00	43.10
XB939	41.50	1.74	5.23	6.39	11.00	54.30
XB940	43.20	1.75	6.64	5.60	29.00	55.00
XB941	46.50	1.70	6.93	5.60	33.00	57.20
XB942	38.90	1.75	6.21	7.52	19.00	53.30
XB943	44.60	1.72	5.98	5.60	29.00	56.10
XB944	22.70	2.03	19.69	15.96	33.00	39.10
XB945	41.80	1.73	5.33	6.73	13.00	54.60
XB946	29.90	1.88	3.11	20.25	8.00	45.90
XB947	42.50	1.74	5.59	6.28	11.00	54.80
XB948	44.20	1.73	5.26	5.82	14.00	55.60
XB949	43.50	1.75	7.91	7.74	33.00	55.00
XB950	28.20	1.86	14.58	13.63	19.00	46.30
Min	21.00	1.61	0.39	4.39	1.00	39.10
Max	58.50	2.03	20.96	24.51	37.00	62.40
Mean	35.49	1.81	7.45	11.68	18.93	50.01
Standard deviation	7.76	0.09	4.59	5.83	9.79	5.35

NB: *ω: water content, F.Angle: Friction angle, and CF: Clay fraction*

5.1.2 PCA computation

The variance of a variable or the covariance between two variables belong to the most important parameters in statistics, particularly in PCA method. The variance-covariance matrix $\mathbf{R}_{6\times6}$ of the standardized matrix $\tilde{\mathbf{X}}_{93\times6}$ compiles the variances as diagonal elements and the covariances as off-diagonal elements. The variance-covariance matrix is calculated as defined by equations 3.2 and 3.4 in chapter 3. Once the data matrix is standardized, the variance-covariance matrix coincides with correlation matrix. An alternative approach of computing the variance-covariance or correlation matrix with the standardized sample matrix is to determine the product of $\tilde{\mathbf{X}}^T\tilde{\mathbf{X}}$.

Table 5.2 shows the linear interrelations among variables as correlation coefficients. The water content is inversely related to density and friction angle with high negative correlation coefficients of -0.92 and -0.84, respectively. In contrast, porosity and water content are proportional with a high positive coefficient of 0.98.

Table 5.2: Linear correlation coefficient matrix between variables.

Variables	ω	Density	Cohesion	F.Angle	CF	Porosity
ω	1.00	-0.92	-0.33	-0.84	0.35	0.98
Density	-0.92	1.00	0.40	0.81	-0.22	-0.96
Cohesion	-0.33	0.40	1.00	-0.03	0.58	-0.34
F.Angle	-0.84	0.81	-0.03	1.00	-0.53	-0.87
CF	0.35	-0.22	0.58	-0.53	1.00	0.32
Porosity	0.98	-0.96	-0.34	-0.87	0.32	1.00

NB: *ω: water content, F.Angle: Friction angle, and CF: Clay fraction*

Table 5.3 shows the eigenvalues, eigenvectors of the correlation matrix and the variance of each variable contributed to the total variance. Some first eigenvectors $\mathbf{V}_r$ define a new set of orthogonal coordinate axes for the best fitting subspace, where the variance of information is maximized. This best fitting subspace will serve as a scaffolding for presentation of original samples as points.

In the case of a two-dimensional best fitting subspace, the first two eigenvectors are taken into account and the best fitting subspace becomes a two-dimensional plane. The firs two eigenvectors alone account for as much as 93% of the total variance in the sample data set. The first eigenvector, which presents the loadings of the first principal component, accounts for more than 65% of the total variance. It weights the contributions of water content and porosity in one direction and density and friction angle with a similar magnitude in the other direction as illustrated in Figure 5.1. The loadings of water content and porosity are present at the above side, while density and friction angle appear at the lower side. The loadings correspond to the positive correlation between water content and porosity that are pointing in the same direction. Density and friction angle, which show an inverse correlation to both water content and porosity, are pointing downwards. Cohesion and clay fraction, which are weakly correlated to water content, density, friction angle, and porosity, are presented with smaller weights in the first eigenvector. Thus, the first principal component reflects the strongest correlation coefficients between the variables as shown in Table 5.2.

Table 5.3: Eigenvectors, eigenvalues and its proportion contributed to variances.

Variables	Eigenvectors					
	I	II	III	IV	V	VI
ω	0.4935	0.0587	-0.1275	-0.2269	0.6874	0.4613
Density	-0.4823	-0.1437	-0.1331	0.4866	0.6579	-0.2436
Cohesion	-0.1334	-0.7096	0.5195	-0.4284	0.1486	-0.0555
F.Angle	-0.4608	0.2060	-0.4623	-0.7101	0.0870	-0.1401
CF	0.2070	-0.6515	-0.6945	0.1011	-0.1996	0.0167
Porosity	0.5009	0.0735	-0.0024	-0.1170	0.1585	-0.8396
	Eigenvalues					
Eigenvalues	3.9180	1.6854	0.2349	0.0908	0.0637	0.0072
Variance proportion (%)	65.30	28.09	3.92	1.51	1.06	0.12
Total variance (%)	65.30	93.39	97.31	98.82	99.88	100

NB: *ω: water content, F.Angle: Friction angle, and CF: Clay fraction*

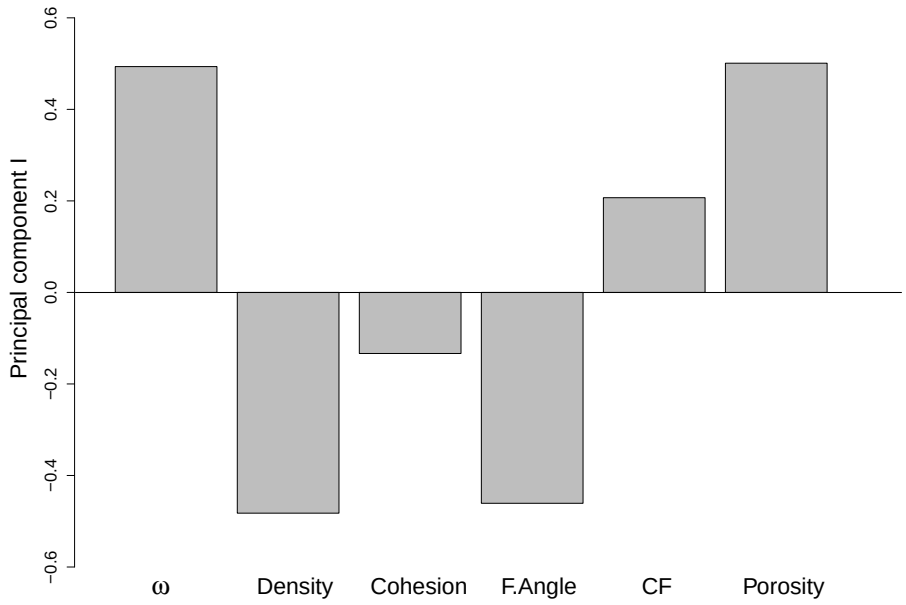


Figure 5.1: Plot of the first principal component loadings.

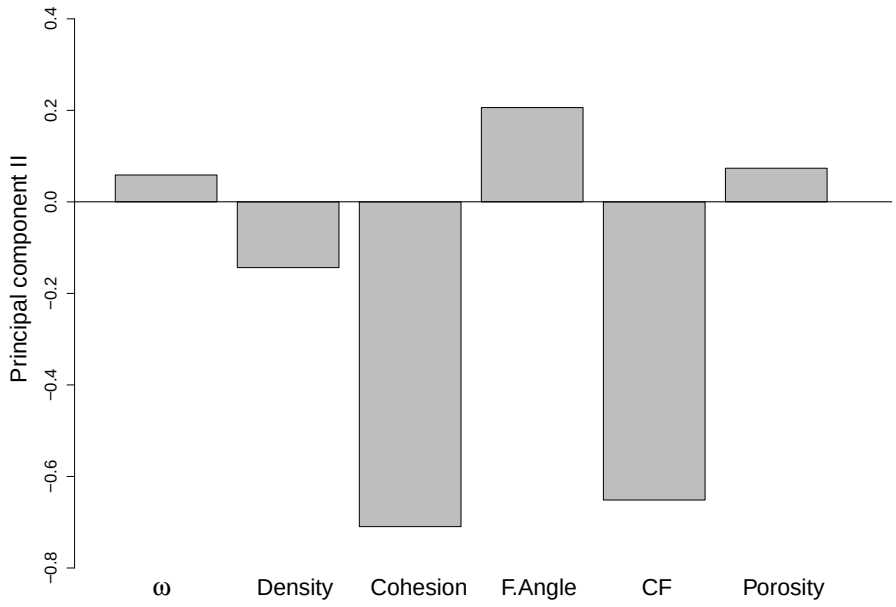


Figure 5.2: Plot of the second principal component loadings.

In a similar manner, Figure 5.2 shows the loadings of the second principal component. Cohesion and clay fraction play the important roles in the second principal component. Since cohesion and clay fraction are positively correlated, they are pointing in the same direction.

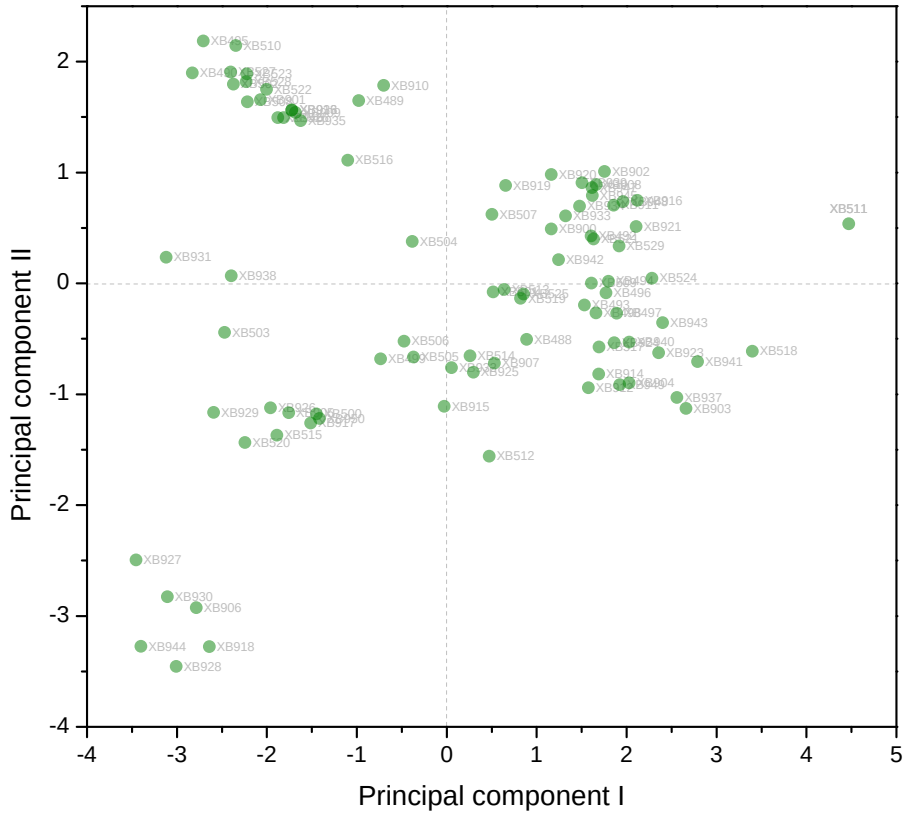


Figure 5.3: Score plot of the first two principal components.

Using PCA method, the data sample matrix is projected into a best fitting subspace. Each property of the soil sample was transformed into a score by projecting it onto the principal axes. The score matrix is a product of standardized matrix $\tilde{\mathbf{X}}_{93\times6}$ and eigenvectors matrix $\mathbf{V}_{6\times6}$ as $\mathbf{Y}_{93\times6} = \tilde{\mathbf{X}}_{93\times6}.\mathbf{V}_{6\times6}$. The variance of variables in the score matrix, tabulated in Table 5.4, are maximized corresponding to eigenvalues and eigenvectors in Table 5.3. In the subspace of the first two principal components, the coordinates of the projection of 93 original samples onto the best fitting plane are plotted in Figure 5.3 with labelled sample names.

Table 5.4: Principal component scores derived from the original data matrix.

Sample	Y1	Y2	Y3	Y4	Y5	Y6
XB949	1.923	-0.914	-0.670	-0.115	0.048	-0.017
XB905	-1.756	-1.165	0.565	-0.637	-0.339	0.032
XB513	0.638	-0.056	0.292	0.385	-0.009	0.031
XB904	2.026	-0.897	-0.393	-0.097	0.167	-0.025
XB906	-2.787	-2.923	-0.213	-0.056	0.168	0.059
XB909	-1.682	1.542	-0.524	-0.167	0.040	-0.078
XB520	-2.245	-1.436	-0.168	0.260	-0.029	0.031
XB948	1.960	0.739	0.549	0.027	0.253	0.031
XB500	-1.451	-1.176	0.625	-0.464	-0.121	-0.126
XB524	2.280	0.046	-0.094	0.086	-0.005	0.039
XB499	-0.735	-0.680	0.568	-0.541	-0.332	-0.219
XB930	-3.108	-2.824	-0.264	0.163	0.318	0.120
XB947	1.613	0.863	0.776	0.031	0.231	0.008
XB950	-1.417	-1.217	0.701	-0.360	-0.160	-0.109
XB931	-3.121	0.236	0.311	0.289	0.156	0.152
XB940	2.026	-0.529	-0.355	0.232	0.030	0.025
XB907	0.528	-0.718	-0.256	0.216	-0.240	-0.022
XB503	-2.470	-0.442	0.073	0.095	-0.850	-0.015
XB495	-2.706	2.186	-0.579	-0.321	0.064	-0.028
XB937	2.560	-1.028	-0.706	-0.288	0.098	-0.015
XB923	2.356	-0.624	-0.622	-0.091	-0.090	-0.008
XB932	0.053	-0.759	-0.124	0.333	0.024	-0.031
XB944	-3.401	-3.272	-0.060	0.273	0.313	0.134
XB517	1.691	-0.573	-0.723	0.475	-0.304	0.037
XB924	1.863	-0.536	-0.397	-0.164	-0.179	-0.057
XB917	-1.516	-1.258	1.237	-0.514	-0.026	-0.126
XB514	0.259	-0.652	-0.223	0.006	0.065	-0.067
XB489	-0.981	1.651	-0.132	-0.399	-0.053	-0.149
XB488	0.886	-0.504	0.194	0.125	-0.649	-0.078
XB508	-2.216	1.639	-0.294	0.122	0.060	0.043
XB515	-1.891	-1.370	-0.950	0.078	0.187	-0.062
XB491	-2.824	1.911	-0.281	-0.202	0.218	-0.066
XB914	1.687	-0.818	-0.553	0.113	0.214	-0.151
XB516	-1.100	1.112	-0.076	0.237	-0.322	-0.006
XB916	2.119	0.749	0.551	-0.119	0.298	0.000
XB521	1.631	0.401	0.258	0.083	0.004	0.018
XB934	1.476	0.697	0.552	0.176	-0.046	0.040
XB903	2.658	-1.129	-0.861	-0.355	0.012	0.029
XB504	-0.383	0.379	-0.848	0.385	-0.062	-0.035
XB938	-2.397	0.069	0.559	0.227	0.038	0.087
XB911	1.854	0.705	0.565	0.009	0.189	0.044
XB928	-3.008	-3.454	-0.214	-0.126	0.103	0.034
XB902	1.753	1.009	0.680	-0.013	0.211	0.017
XB925	0.297	-0.802	-0.098	0.004	-0.335	0.178
XB512	0.472	-1.558	-0.083	-0.108	0.140	-0.057
XB942	1.242	0.215	0.225	0.097	-0.183	-0.022
XB490	-2.831	1.898	-0.077	-0.431	0.097	-0.018
XB498	1.659	-0.265	0.388	0.090	-0.291	-0.036
XB921	2.105	0.515	0.374	-0.058	0.165	0.027
XB507	0.501	0.624	-0.145	1.036	-0.517	0.069
XB526	-1.814	1.495	-0.334	0.033	0.009	-0.039
XB910	-0.703	1.786	-0.047	-1.007	-1.016	0.444
XB509	1.607	0.005	0.691	0.128	0.304	0.040
XB528	-2.232	1.818	-0.185	0.127	0.103	0.034
XB510	-2.345	2.147	-0.583	0.163	0.317	0.009
XB945	1.618	0.795	0.595	-0.009	0.047	0.021
XB936	-1.723	1.564	-0.287	-0.219	0.099	-0.083
XB915	-0.030	-1.107	-0.122	0.346	0.054	-0.018
XB522	-2.006	1.750	-0.189	-0.013	-0.013	-0.040
XB519	0.820	-0.133	-0.256	0.130	-0.358	-0.051
XB912	1.572	-0.940	-0.243	0.057	-0.092	-0.036
XB511	4.470	0.539	-0.074	-0.898	0.615	0.196
XB493	1.528	-0.195	-0.020	0.204	-0.205	-0.062
XB523	-2.224	1.891	-0.227	0.065	-0.010	0.032
XB497	1.887	-0.268	-0.162	0.040	-0.205	-0.053
XB933	1.319	0.609	0.366	0.329	0.272	-0.178
XB943	2.400	-0.353	-0.408	0.064	-0.057	0.026
XB939	1.504	0.909	0.743	0.091	0.118	0.028
XB926	-1.961	-1.123	0.669	-0.314	0.130	-0.043
XB918	-2.641	-3.277	0.155	-0.428	-0.004	-0.044
XB908	1.663	0.893	0.651	0.154	0.176	0.058
XB946	-1.878	1.497	-0.399	-0.142	0.078	-0.040
XB505	-0.371	-0.663	0.570	-0.140	-0.322	-0.121
XB506	-0.477	-0.522	-0.061	0.495	-0.410	-0.008
XB518	3.396	-0.612	-1.222	-0.558	0.247	0.076
XB919	0.652	0.884	0.652	0.099	0.150	0.035
XB927	-3.456	-2.494	0.326	0.173	0.415	0.111
XB920	1.161	0.983	0.778	0.212	-0.028	0.031
XB941	2.789	-0.704	-0.586	-0.173	-0.055	0.016
XB913	-1.723	1.564	-0.287	-0.219	0.099	-0.083
XB525	0.853	-0.097	-0.024	0.373	-0.095	0.008
XB922	2.560	-1.028	-0.706	-0.288	0.098	-0.015
XB929	-2.594	-1.162	0.862	0.197	0.227	0.101
XB935	-1.626	1.468	-0.321	-0.050	0.205	-0.037
XB501	0.516	-0.077	-0.057	0.560	0.215	0.043
XB529	1.914	0.337	0.387	0.077	0.055	0.014
XB527	-2.404	1.908	-0.305	0.061	0.016	0.021
XB496	1.770	-0.086	0.217	-0.083	-0.255	-0.056
XB502	-2.373	1.798	-0.342	0.039	0.011	-0.013
XB494	1.796	0.019	0.246	0.019	-0.105	-0.039
XB492	1.605	0.427	0.452	0.058	0.232	0.051
XB901	-2.072	1.657	-0.308	0.092	0.182	0.023
XB900	1.161	0.491	0.337	0.125	-0.008	-0.018
Mean	0.000	0.000	0.000	0.000	0.000	0.000
Total variance	3.918	1.685	0.235	0.091	0.064	0.007
Variance proportion (%)	65.30	28.09	3.92	1.51	1.06	0.12
Cummulative total variance (%)	65.30	93.39	97.31	98.82	99.88	100

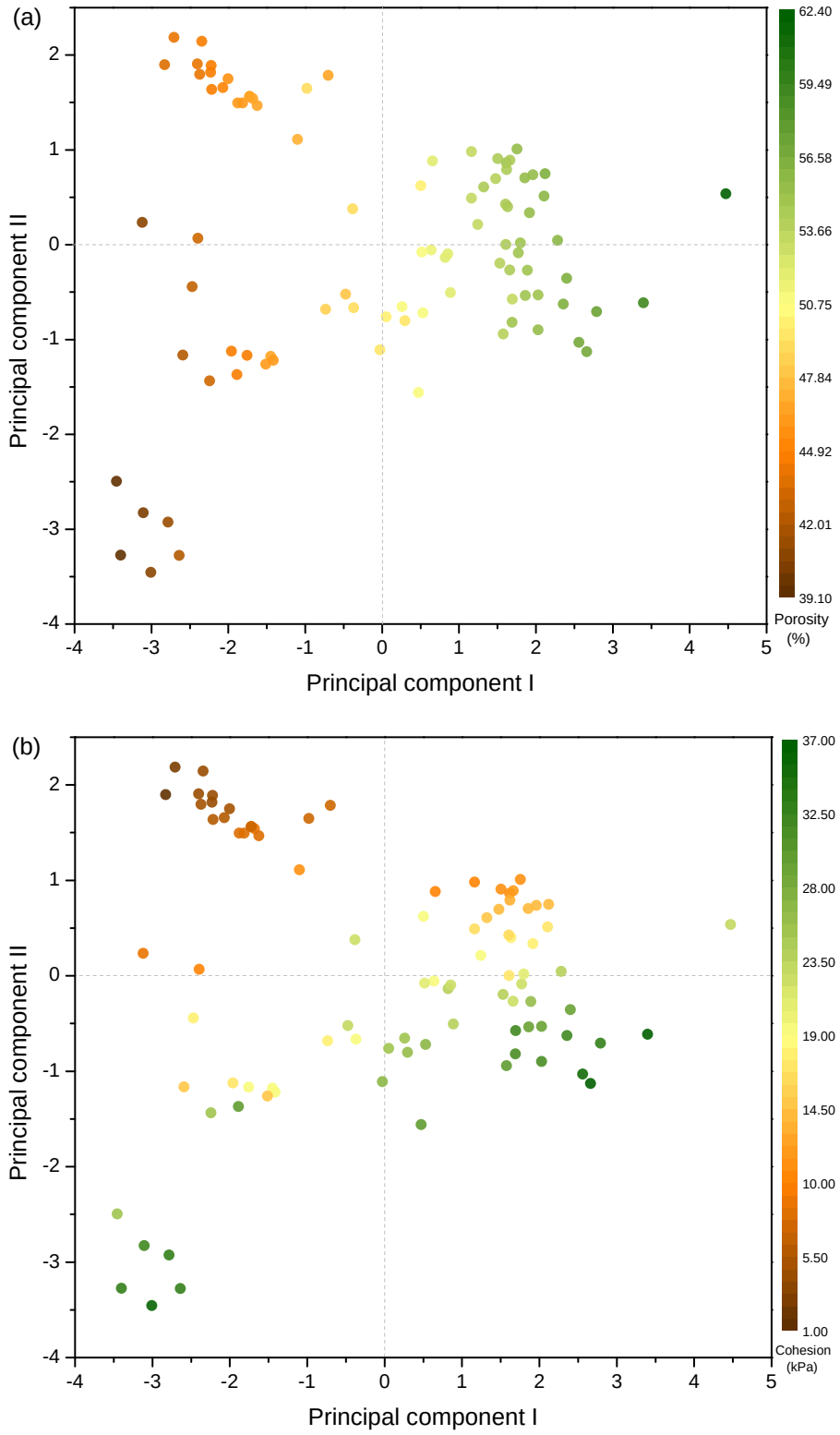


Figure 5.4: PCA plot of the first two principal components with reference to: (a) Porosity, and (b) Cohesion.

As mention above, the loading on the first principal component shows the largest proportion of porosity variable contributed to the total variance. The PCA plot in Figure 5.4(a) with reference to porosity shows clearly the trend of porosity along the first principal component. Porosity increases gradually from the left (negative) side in brown color towards the right (positive) side in dark green color as predicted by the positive loading.

On the second principal component, cohesion has largest proportion contributed to the variance. It points to the negative side as shown in Figure 5.2. Cohesion in PCA plot Figure 5.4(b) also presents this trend in vertical direction of the second principal component.

The distribution and the importance of variables in variance contribution of the original samples can be evaluated by the plots of loadings and PCA. This becomes more clear when the variable axes are added over the PCA plot as shown in the next section.

5.1.3 PCA visualization

In PCA, the score matrix $\mathbf{Y}$ is a transformation of the original sample data matrix as mentioned above. It is the projection of original sample in to a lower dimensional subspace. The visualization of the soil samples with various properties in a global view requires plotting both soil samples as points and their properties as axes in two- or three-dimensional biplots. The original samples are projected on the plane of the two first principal components as scores in Figure 5.3. In addition, variables are overlaid on as axes as shown in Figure 5.5. The cosine of the angle between two axes exhibits the correlation coefficient between corresponding variables.

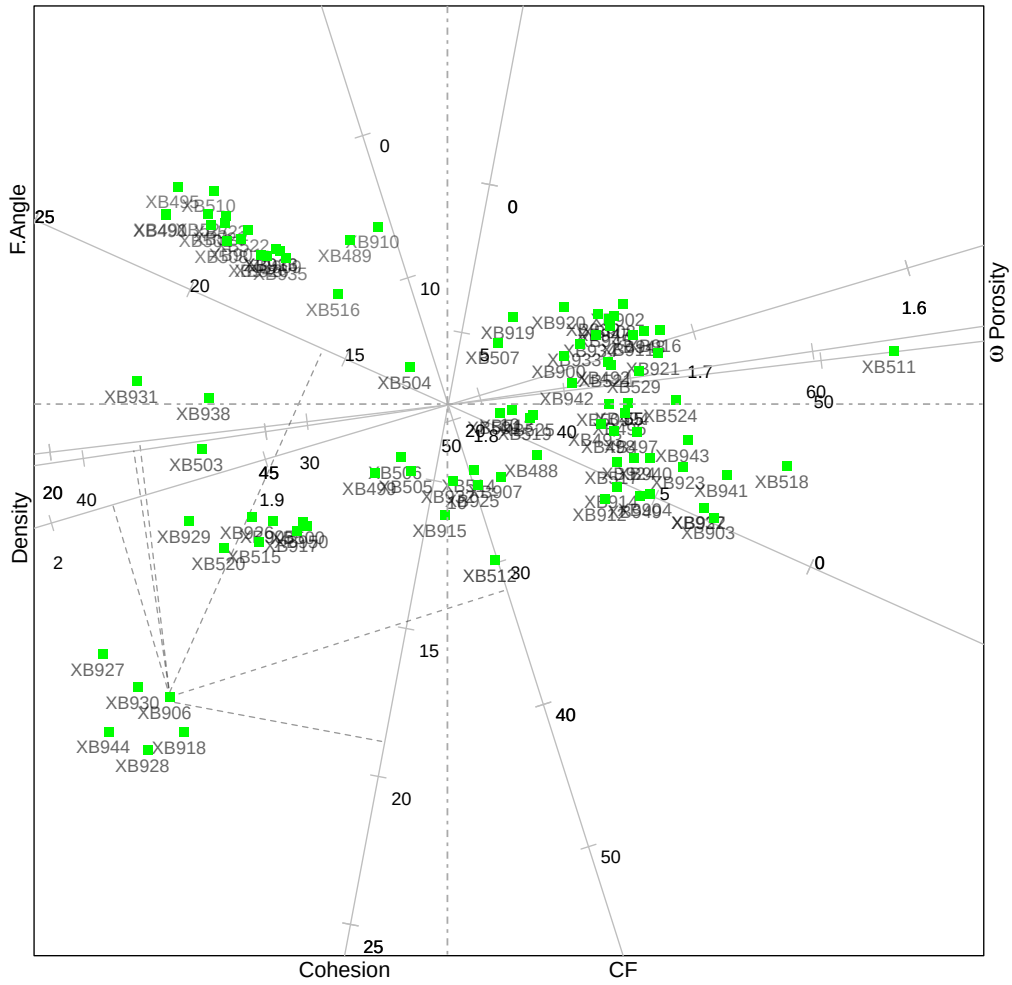


Figure 5.5: PCA biplots of the first two principal components.

Table 5.5: Adequacies and predictivities of variables on the two-dimensional subspace.

Variables	Adequacy	Predictivity	
		Component I	Component I&II
ω	0.25	0.95	0.96
Density	0.25	0.91	0.95
Cohesion	0.52	0.07	0.92
F.Angle	0.25	0.83	0.90
CF	0.47	0.17	0.88
Porosity	0.26	0.98	0.99

NB: *ω: water content, F.Angle: Friction angle, and CF: Clay fraction*

The principal axes provide an essential scaffolding, on which the biplots are based. But in visualization, the scaffolding values are normally not considered. The

presentations of samples as points and variables as axes and the quality of their display get more attentions. The overall quality of the PCA approximation is 0.93, implying that 93% of the variation in the samples is accounted for the first two principal components.

Besides overall quality, there is also interest in the quality of display of the variables in r dimensions. Gardner-Lubbe et al. (2008) proposed the measures of fit of variables as adequacy and predictivity. Table 5.5 compiles the measures of fit of the approximating variables on a two-dimensional plane. Adequacy gets maximum of unity when at $r = 6$ the transformation is exact. For high adequacy, the variable lies in or near the r-dimensional subspace, and for low adequacy the corresponding variable lies nearly orthogonal to the subspace. In a two-dimensional display, the adequacy of water content gets its maximum in the plane of components I and component V. The maximum of adequacy of cohesion is displayed in the plane of component II and component III and so on as shown in Table 5.3. While adequacy is associated with the visualization of variables, axis predictivity is the variance accounted for each variable. The axis adequacies and predictivities of variables are compiled in Table 5.5. Cohesion and clay fraction with the highest adequacies appear closest to the display plane. In a two-dimensional plane of component I and component II, all variables have quite high predictivities. The lowest predictivity of 0.88 of the clay fraction means that 88% of information is displayed in the plane of the first two principal components.

Table 5.6: Relative errors of variables of sample XB 906.

Variable	ω	Density	Cohesion	F. Angle	CF	Porosity
	(%)	(g/cm^3)	(kPa)	(deg)	(%)	(%)
Predicted values	23.48	1.97	18.68	15.65	31.93	41.37
Actual values	24.90	1.98	18.38	16.49	33.00	41.00
(Max - Min) range	37.50	0.42	20.57	20.12	36.00	23.30
Relative error (%)	3.79	2.14	1.47	4.16	2.98	0.40

NB: *ω: water content, F.Angle: Friction angle, and CF: Clay fraction*

Another measure of the goodness of the approximation is the relative error, which can be computed for any variable of any sample. The relative error is defined as the difference between the predicted and actual values, expressed as a percentage of the range between maximum and minimum of the actual values of the particular variable. Table 5.6 compiles the relative errors of all variables of the arbitrary sample XB 906.

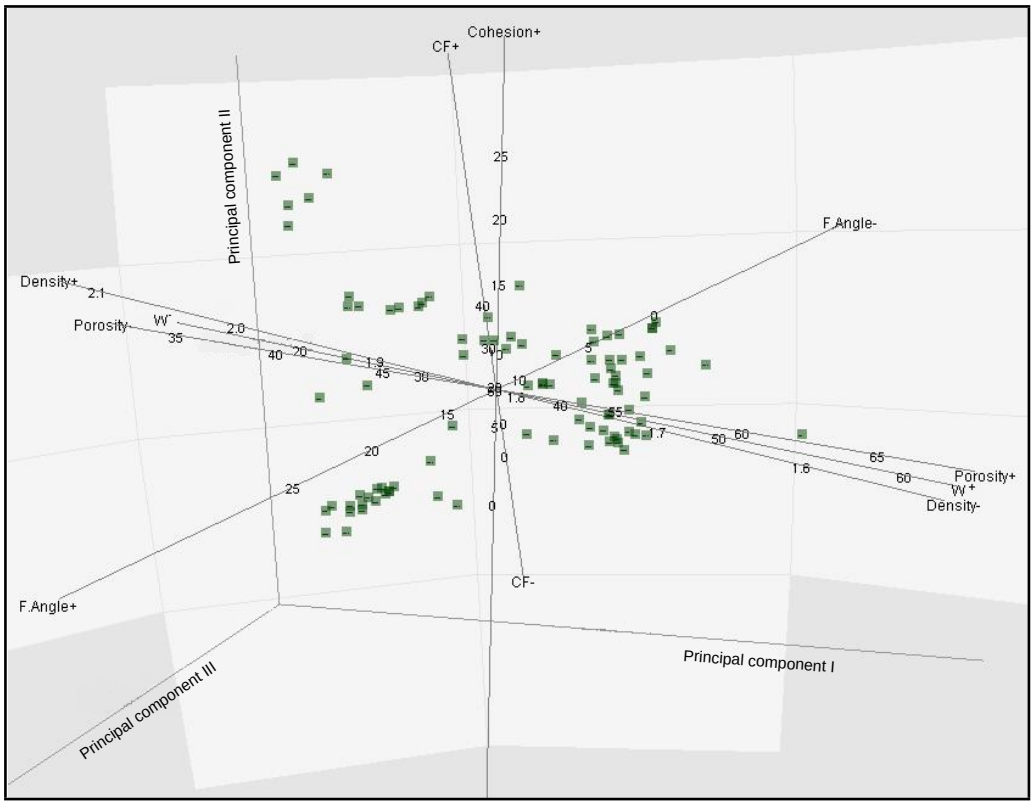


Figure 5.6: PCA biplots on the first three principal components.

Figure 5.6 presents a PCA biplots of the first three principal components. The overall quality of the three-dimensional display is 0.97 that implies that more than 97% of data information is displayed. It becomes evident that water content, density and porosity point significantly in the direction of the first component, cohesion and clay fraction play an important role of the second component while the friction angle heavily weights the third component.

5.1.4 Cluster analysis

In PCA method, the interaction between the variables is investigated in detail. The columns of the sample matrix are under consideration. In cluster analysis, the samples are the main objects of interest. Specifically, the distances between pairs of samples are calculated, compared and on the base of the nearest similar distances a dendogram is built up. The original soil sample matrix $\mathbf{X}_{93\times6}$ as compiled in Table 5.1 is investigated. The distances between each pair at samples are calculated as formulated in equation 4.1. As in PCA, the original sample matrix should be standardized prior to calculating the distance.

The number of pair distances corresponding to all the combination of two samples out of n = 93 soil samples results in n(n-1)/2 = 4,278 distances. The computation of the differences between all possible pairs of samples results in a square symmetrical matrix $\mathbf{D}_{93\times93}$ where each element is the distance of the two samples indicated by the corresponding column and row.

The samples with the highest similarities are arranged and placed into a hierarchy, then those pairs are merged and the matrix is recomputed. The process iterates until the dissimilarity matrix reaches the rank of two. The lower distance indicates that the soil samples are similar and a group or cluster is created. A certain group is more or less homogeneous in distances and distinct from the others. The constructed dendrogram is presented in the tree form as shown in Figure 5.7. The height of the tree between two sample or two groups of sample exhibits the distances between those samples or groups of sample. According to the dendrogram, the soil samples of similarity in distance are clustered into 8 groups. The soil sample of organic clay, XB 511 is out all group. The resulting groups of soils should be similar with those from PCA biplot as shown in the next section.

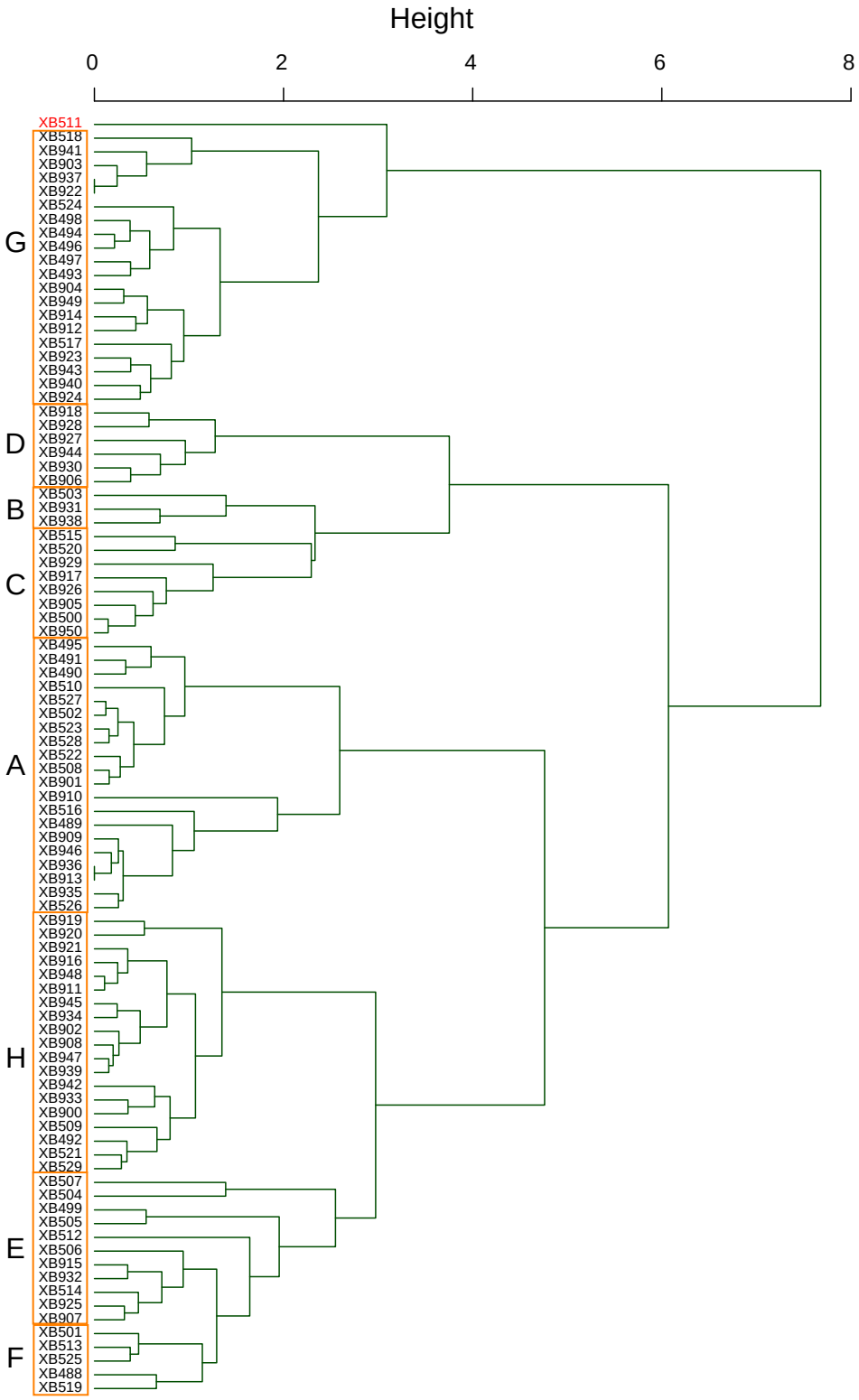


Figure 5.7: Dendrogram of soil sample matrix clustered by linkage method and soil groups associated with PCA biplot.

5.1.5 Combination of cluster analysis and PCA

The groups of soil can be clustered according to the PCA biplot. The soil groups sketched in elliptic shapes with the aid of the dendrogram from cluster analysis are presented in Figure 5.8.

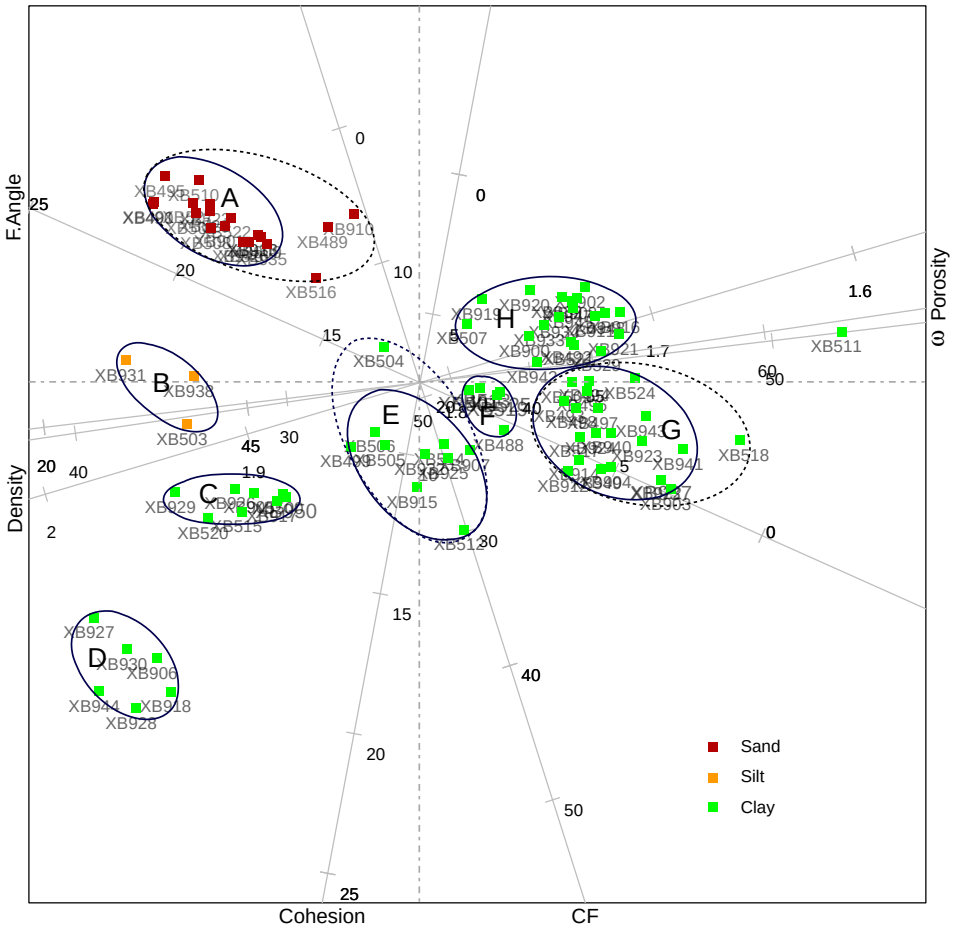


Figure 5.8: Clustered soil groups a PCA biplot of the first two principal components.

The soil group A in red color, which is displayed in the second quarter, represents sands with low clay fraction and cohesion, while density and friction angle are high. Three soil samples XB 516, XB 489 and XB 910 are clustered in group A in term of distance form the dendrogram, though they are classified as clayey and silty sand, respectively. The soil samples in group B represent silts. The samples in group C

are classified as clays. The clayey soils from greater depth of 25 m to 35 m appear in group D with the values of density, clay fraction, and cohesion being very high. The soil samples in groups E, F, G, H represent clays from moderate depth of 5 m to 20 m. The two samples XB 511 and XB 518 are classified as organic clay (OH) with extreme values of water content and porosity. The sample XB 518 can be grouped in group G, but the sample XB 511 can not be clustered in any group. The silty clay sample XB 504 with lower clay fraction which originates from shallow depth can be clustered into group E in accordance with cluster analysis.

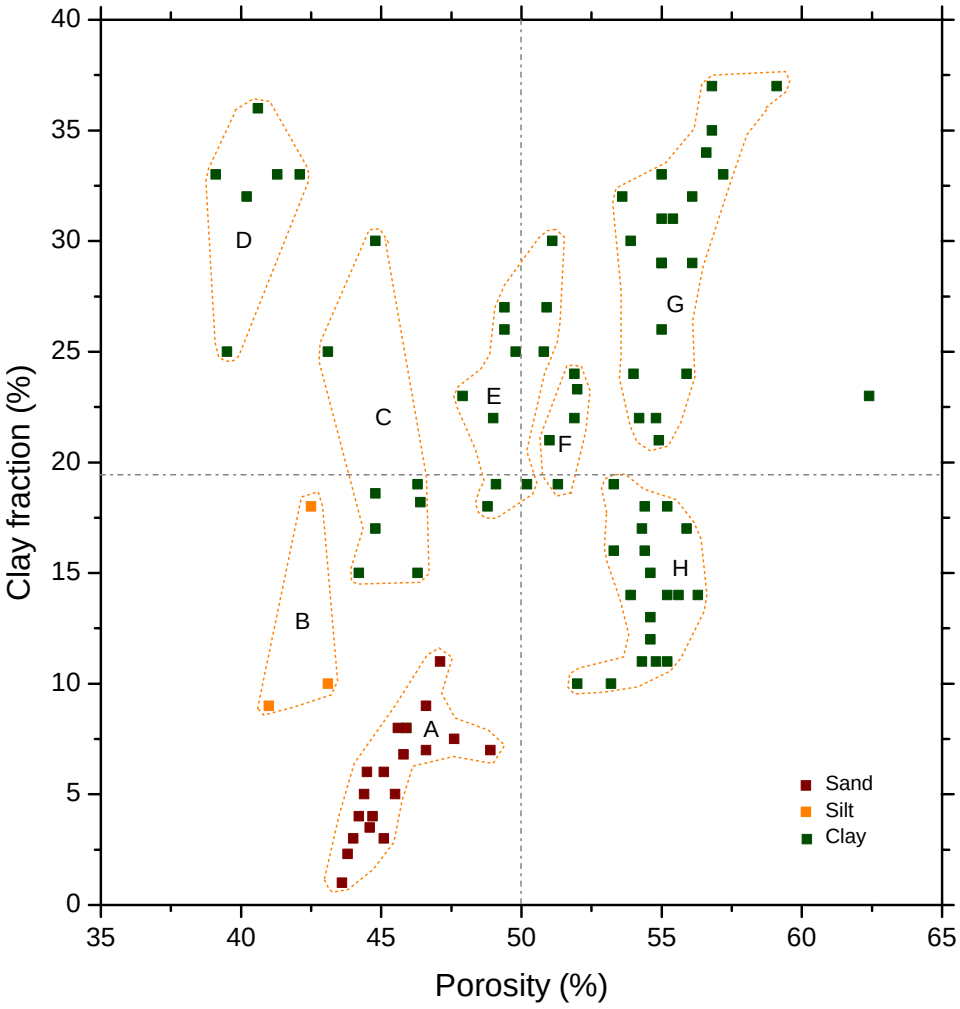


Figure 5.9: Soil groups in the cross-plot of porosity and clay fraction.

In the PCA biplot, soil groups are clustered and viewed globally by six properties of soil samples in one graph. Figure 5.9 shows the cross-plot of porosity and clay fraction of the soil samples. These properties play important roles in the first and second principal component. The eight groups of soils are sketched in a similar way to the PCA biplot. Though the samples in the individual groups are slightly more scattered, the cross-plot of porosity and clay fraction clearly indicates the different groups of soils.

5.2 Example 2: Petrophysical and geotechnical data

5.2.1 Data preparation

The petrophysical and geotechnical properties of 77 soil samples collected from the dike system in Germany are described by six parameters: volumetric water content, real part of relative dielectric permittivity, density, porosity, resistivity and magnetic susceptibility as shown in Table 5.7. Because of the wide variation of the values of resistivity and magnetic susceptibility the logarithms of the values are regarded. The data set is prepared as a matrix of 77 rows and six columns in a random manner. The rows indicate the samples, while the columns describe their geotechnical and petrophysical properties. Since the variables are measured in different dimensions, it becomes necessary to standardize the values of each column by centring and normalizing to standard deviation.

Table 5.7: Petrophysical and geotechnical properties of soil samples from Germany.

Sample	ω*	R.P200	Density	Porosity	Log.Res.	Log.M.S.
	(%)	-	(g/cm^3)	(%)	(Ωm)	(10^{-8} m^3/kg)
XB 450	0.32	16.38	1.88	0.40	1.97	1.26
XB 451	0.36	16.46	1.85	0.47	1.77	1.01
XB 452	0.27	12.79	1.78	0.42	1.86	1.25
XB 453	0.18	9.23	1.92	0.38	1.75	1.13
XB 454	0.23	11.06	1.85	0.41	1.73	1.27
XB 456	0.24	11	2.04	0.41	2.36	0.64
XB 457	0.21	10.87	1.83	0.38	1.95	1.25
XB 458	0.23	10.23	1.8	0.45	2.71	0.58
XB 459	0.4	22.84	1.78	0.51	1.90	1.19
XB 460	0.19	7.14	1.1	0.68	1.85	1.90
XB 461	0.02	3.14	1.19	0.50	2.12	0.88
XB 463	0.09	5.06	1.5	0.54	1.67	1.43
XB 464	0.14	6.95	1.36	0.52	1.64	1.12
XB 465	0.38	20.33	1.83	0.49	1.53	1.04
XB 467	0.03	4.07	1.54	0.35	2.48	0.74
XB 468	0.2	9.15	1.91	0.41	2.35	0.71
XB 469	0.06	4.12	1.42	0.56	1.44	1.69
XB 470	0.12	6.42	1.36	0.53	1.50	1.23
XB 472	0.08	6.25	1.43	0.47	1.70	1.06
XB 473	0.35	20.51	1.81	0.41	1.82	0.95
XB 476	0.06	4.01	1.45	0.52	2.03	1.02
XB 477	0.06	5.48	1.44	0.48	1.58	1.58
XB 479	0.01	4.19	1.35	0.44	1.51	1.43
XB 484	0.21	8.91	1.32	0.44	1.33	1.88
XB 488	0.05	3.62	1.21	0.55	1.95	0.96
XB 700	0.37	20.34	1.68	0.44	1.28	1.08
XB 701	0.12	8.75	1.93	0.33	1.67	1.07
XB 702	0.24	11.84	1.62	0.41	2.19	0.63
XB 703	0.28	13.93	1.84	0.40	2.22	0.44
XB 704	0.24	11.49	1.58	0.51	1.22	1.59
XB 705	0.1	5.65	1.48	0.52	1.74	1.05
XB 706	0.17	8.21	1.83	0.40	1.25	0.72
XB 707	0.12	6.34	1.55	0.51	2.37	1.12
XB 708	0.05	5.14	1.71	0.41	1.40	1.39
XB 709	0.06	6.02	1.59	0.36	1.53	1.17
XB 710	0.32	19.47	1.96	0.36	1.92	0.82
XB 711	0.16	8.11	1.46	0.52	1.51	1.31
XB 712	0.13	6.9	1.8	0.38	2.25	0.68
XB 713	0.18	8.38	1.35	0.48	1.26	1.29
XB 714	0.36	18.61	1.89	0.43	1.37	1.54
XB 715	0.32	13.5	1.8	0.40	1.21	1.23
XB 716	0.15	7.02	1.69	0.50	1.34	0.88
XB 717	0.02	3.24	1.18	0.54	1.67	1.11
XB 718	0.37	16.08	1.8	0.45	1.66	1.18
XB 719	0.04	3.83	1.37	0.49	2.00	0.93
XB 720	0.31	17.7	1.94	0.38	1.38	1.00
XB 721	0.08	5.03	1.53	0.56	1.42	1.15
XB 722	0.16	8.05	1.67	0.40	2.13	0.50
XB 723	0.13	7.27	1.8	0.43	1.35	1.43
XB 724	0.06	4.36	1.19	0.50	1.25	1.45
XB 725	0.24	13.62	1.94	0.37	1.63	1.18
XB 726	0.07	5.16	1.38	0.44	1.42	1.26
XB 727	0.23	12.49	1.88	0.42	1.30	1.07
XB 728	0.26	14.7	2.08	0.36	1.49	1.06
XB 729	0.12	6.65	1.52	0.33	1.75	1.55
XB 730	0.18	9.06	1.87	0.43	1.33	1.08
XB 731	0.13	6.24	1.6	0.46	2.29	0.57
XB 732	0.06	3.75	1.58	0.52	1.77	0.99
XB 733	0.15	7.41	1.67	0.51	1.31	1.23
XB 734	0.32	14.51	1.9	0.46	1.49	1.14
XB 735	0.16	8.35	1.46	0.49	1.20	1.32
XB 736	0.01	3.17	1.73	0.40	1.82	1.09
XB 737	0.29	14.18	1.63	0.42	1.90	1.04
XB 738	0.06	4.38	1.49	0.45	1.59	1.00
XB 739	0.01	2.81	1.54	0.45	1.88	0.97
XB 740	0.02	4.33	1.6	0.42	1.11	1.55
XB 741	0.13	8.64	1.5	0.46	1.23	1.06
XB 744	0.03	7.11	1.65	0.45	1.39	1.51
XB 747	0.16	8.3	1.36	0.43	1.42	1.08
XB 748	0.27	15.43	1.98	0.34	1.68	1.19
XB 749	0.25	12.37	1.8	0.39	2.12	1.17
XB 750	0.12	6.79	1.6	0.49	1.43	1.23
XB 751	0.36	20.82	1.98	0.41	1.56	1.41
XB 752	0.17	9.6	1.57	0.44	1.53	1.18
XB 753	0.28	16.2	1.8	0.37	1.40	1.13
XB 754	0.34	15.87	1.85	0.45	1.36	1.31
XB 755	0.17	9.4	1.37	0.46	1.64	1.19
Min	0.01	3.14	1.10	0.33	1.22	0.44
Max	0.40	22.84	2.04	0.68	2.71	1.90
Mean	0.18	9.67	1.63	0.45	1.82	1.11
Standard deviation	0.11	5.53	0.24	0.07	0.37	0.35

NB. *: ω*: volumetric water content,*

R.P200: Real part of dielectric permittivity at 200 MHz,

Log.Res.: Logarithms of resistivity,

Log.M.S.: Logarithms of magnetic susceptibility.

5.2.2 PCA computation and visualization

The matrix of correlation coefficients between variables is presented in Table 5.8. Volumetric water content shows a very high positive correlation +0.96 with the real part of relative dielectric permittivity, a lower positive correlation with density. Density and porosity are inversely related with the correlation coefficient of -0.66. Resistivity and magnetic susceptibility are also found to be inversely correlated. It can be assumed that the soils with a higher iron content are characterized by lower electrical resistivity or higher conductivity.

Figure 5.10 displays the correlations between variables and the degree of approxi-

Table 5.8: Correlation coefficients between variables.

Variable	ω^*	R.P200	Density	Porosity	Log.Res.	Log.M.S.
ω^*	1.00	0.95	0.64	-0.65	-0.02	-0.07
R.P200	0.95	1.00	0.68	-0.68	-0.04	-0.09
Density	0.64	0.68	1.00	-1.00	0.09	-0.29
Porosity	-0.65	-0.68	-1.00	1.00	-0.08	0.27
Log.Res.	-0.02	-0.04	0.09	-0.08	1.00	-0.60
Log.M.S.	-0.07	-0.09	-0.29	0.27	-0.60	1.00

NB. *: ω *: volumetric water content, R.P200: Real permittivity at 200 Hz, Log.Res.: Logarithms of resistivity, Log.M.S.: Logarithms of magnetic susceptibility .*

mation of each. Correlations between variables are indicated by the cosine of the angle between them. The degree of approximation of each variable is indicated in the brackets or by the lengths of the arrows as the unit correlation of exact representations is given by the square root of $\text{diag}(\mathbf{V}_r\mathbf{\Sigma^2}\mathbf{V}_r^T)$ (Gower et al., 2010). Eigenvectors and eigenvalues of the correlation matrix are compiled in Table 5.9. The contribution of eigenvalues and corresponding eigenvectors to the total variance are also presented in the table. The first eigenvalue with the highest magnitude accounts for as much as 48% to the total variance. The first two eigenvalues contribute as much as 75% to the total variance.

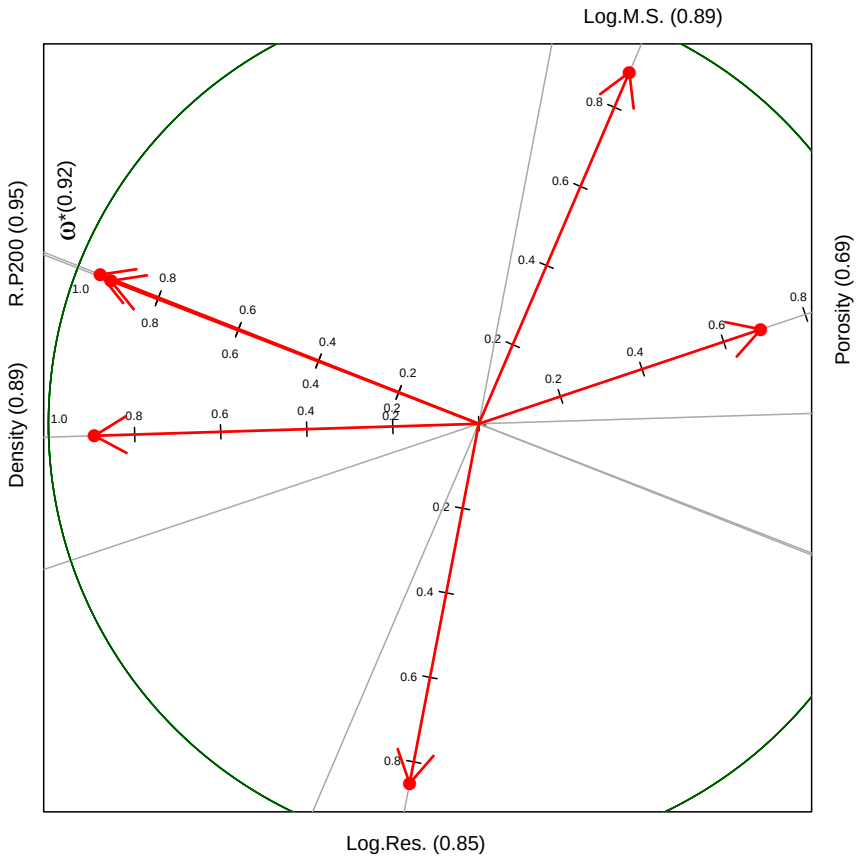


Figure 5.10: Correlation and axis approximation of variables.

Table 5.9: Eigenvectors, eigenvalues and its proportion contributed to variances

Variables	Eigenvectors					
	I	II	III	IV	V	VI
ω*	-0.5038	0.2589	-0.4016	0.0080	-0.1717	0.6988
R.P200	-0.5182	0.2703	-0.2976	0.0159	-0.2594	-0.7087
Density	-0.5263	-0.0217	0.2524	-0.0331	0.8105	-0.0269
Porosity	0.3862	0.1711	-0.7321	-0.2375	0.4707	-0.0873
Log.Res.	-0.0948	-0.6516	-0.3701	0.6516	0.0619	-0.0318
Log.M.S.	0.2061	0.6368	0.1162	0.7194	0.1443	0.0067
	Eigenvalues					
Eigenvalues	2.8789	1.6444	0.8418	0.3762	0.2247	0.0339
Variance proportion(%)	47.98	27.41	14.03	6.27	3.75	0.57
Cummulative total variance(%)	47.98	75.39	89.42	95.69	99.43	100

NB. *: ω*: volumetric water content, R.P200: Real part of dielectric permittivity at 200 MHz, Log.Res.: Logarithms of resistivity, Log.M.S.: Logarithms of magnetic susceptibility.*

Figure 5.11 and 5.12 present the first and second principal component loadings, respectively. The loadings of the first principal component show a high proportion

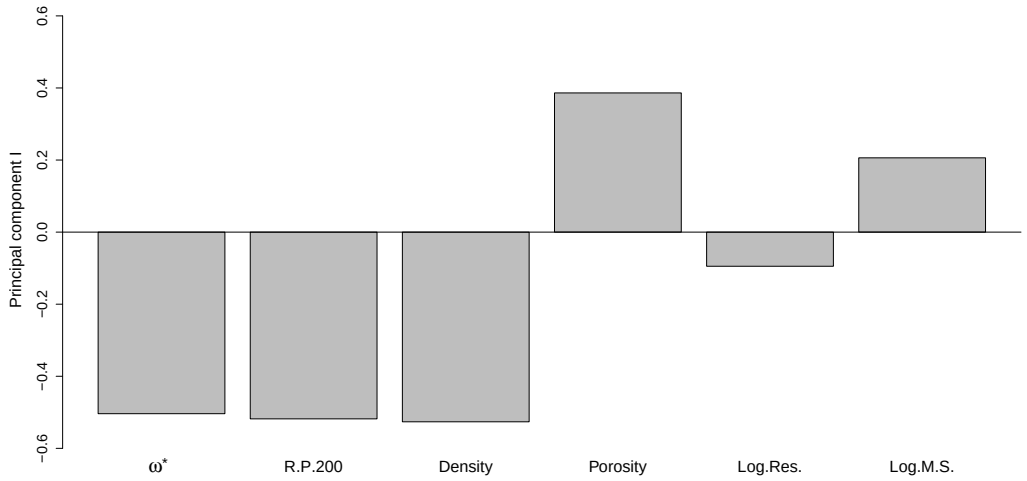


Figure 5.11: Plot of the first principal component loading.

contributed by porosity at the positive side and of water content, real part of relative dielectric permittivity and density at the negative side. The distribution of the loadings reflects the strong correlation between the first four variables. The loadings on the second principal component exhibit a heavy weight of resistivity and magnetic susceptibility in reverse directions, which reflects the negative correlation between the two variables.

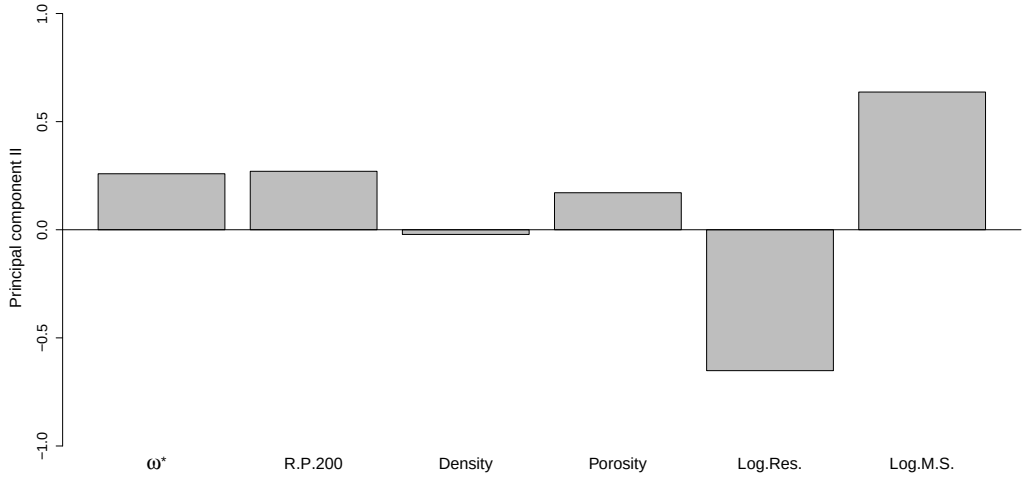


Figure 5.12: Plot of the second principal component loading.

The original soil sample matrix can be approximated by a projection into the best

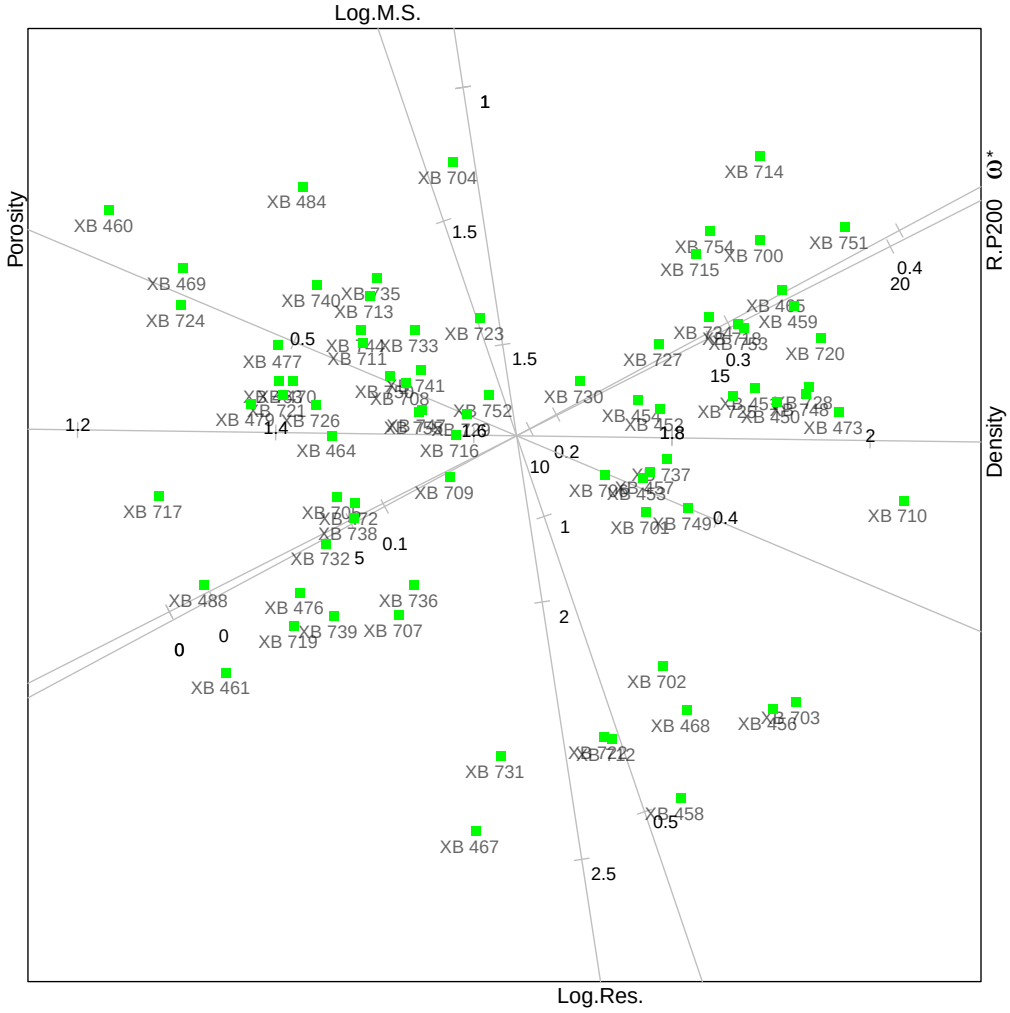


Figure 5.13: PCA biplots of the first two principal components.

fitting subspace of lower dimension. In a two-dimensional subspace, each original observation was converted by a projection of the principal component scores into the best fitting plane spanned by the first two principal component axes as shown in Figure 5.13. The variables are superposed by the display of the corresponding parameter functions as unit interpolants. The overall quality of the display is 0.75 denoting that 75% of the total variance is presented by the PCA display of the first two principal components. The adequacy and predictivity of variable axes of the first two principal components are summarized in Table 5.10. Resistivity and magnetic susceptibility are characterized by high values of adequacy denoting that their axes

Table 5.10: Adequacies and predictivities of variables on a two-dimensional subspace.

Variables	Adequacy	Predictivity	
		Component I	Component I & II
ω*	0.32	0.73	0.84
R.P200	0.34	0.77	0.89
Density	0.28	0.80	0.80
Porosity	0.18	0.43	0.48
Log.Res.	0.43	0.03	0.72
Log.M.S.	0.45	0.12	0.79

NB. *: ω*: volumetric water content,*
R.P200: Real part of dielectric permittivity at 200 MHz,
Log.Res.: Logarithms of resistivity,
Log.M.S.: Logarithms of magnetic susceptibility.

appear close to the plane of the two first principal components. The predictivity of all variables considering the principal components I and II varies between 0.72 and 0.89 except the predictivity of porosity with the low value of 0.48. The poor predictivity of porosity indicates that the porosity axis makes the largest angle with the plane of the principal component I and II. The greatest part of variance is lost compared to the other variables as visible in Table 5.9 and Figure 5.14.

On the PCA display, properties of every sample point can be predicted on any variable by a projection at the corresponding axis. The cosines of the angles between the variables represent their linear correlations. The soil samples that are presented as point scatters in the PCA biplot indicate trends in the distribution of the investigated properties.

Figure 5.14 shows a PCA biplot of the first three principal components. As much as 89 % of the total variance of the data matrix is displayed in the three-dimensional subspace. Volumetric water content, real part of relative dielectric permittivity and density play an important role of the first principal component. Resistivity and magnetic susceptibility indicate significantly the direction of the second principal component and the porosity heavily weights the third principal component.

5.2.3 Cluster analysis

The petrophysical and geotechnical properties of 77 soil samples from Germany, described in Table 5.7, are compiled as a matrix of 77 rows and 6 columns. The rows are the soil samples and columns are their properties. In cluster analysis, the similarity in term of distance between samples is of interest. The distances between each pair of samples are considered. The combinations of two samples out of n = 77 samples result in n(n-1)/2 = 2,926 distances.

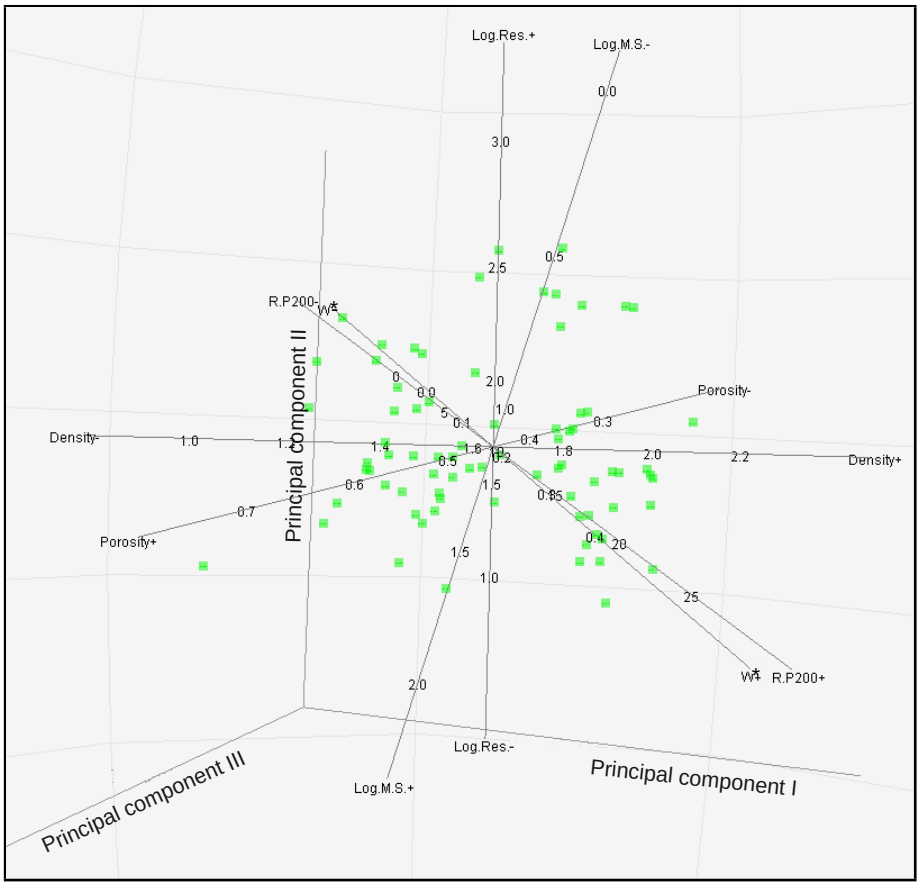


Figure 5.14: PCA biplot of the first three principal components.

From the matrix of distances, the dendrogram is constructed with the rule that the highest similarities in distance are grouped in a cluster. The resulting dendrogram is shown in Figure 5.15.

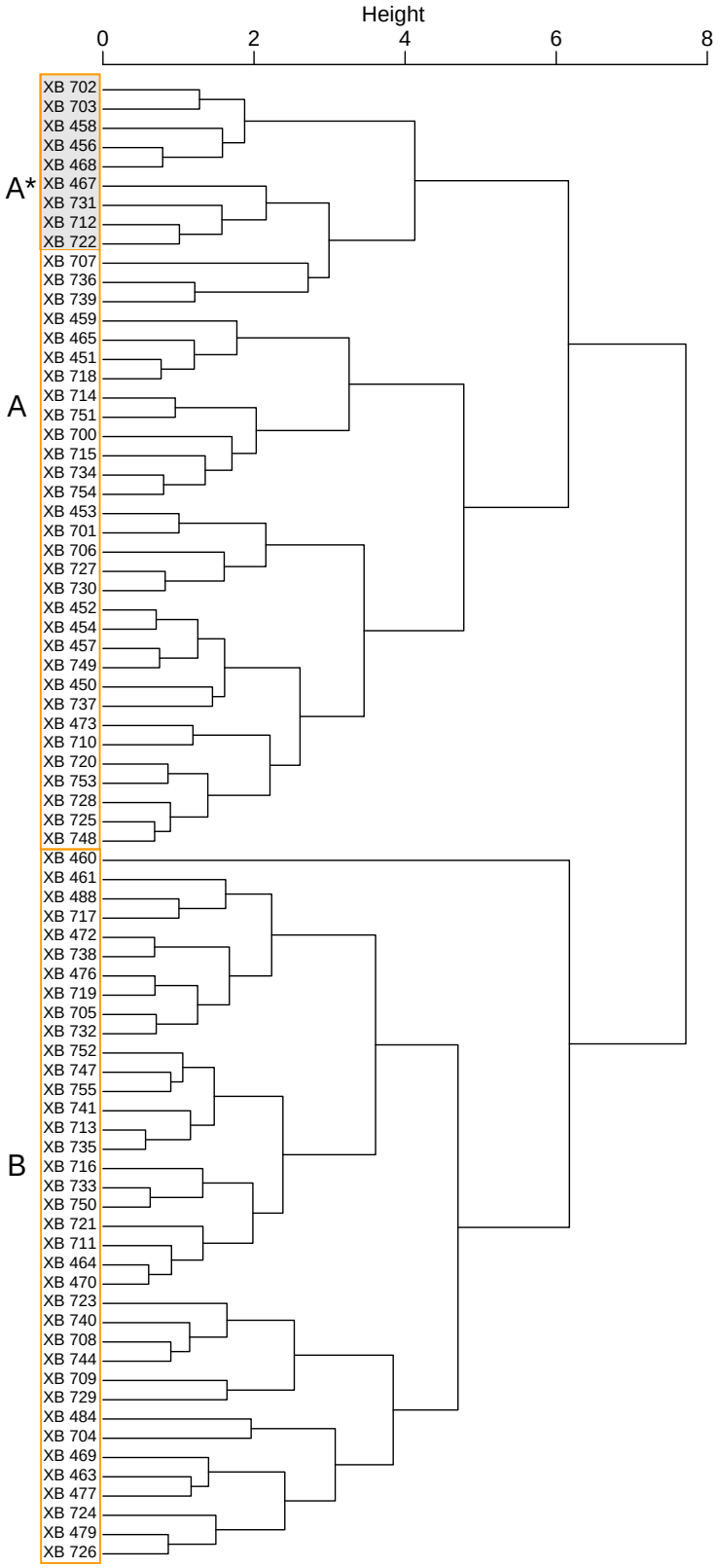


Figure 5.15: Dendrogram of soil sample matrix clustered by linkage method and soil groups associated with PCA biplot.

In cluster analysis, the clusters are defined by the distances between the pairs of samples. The height of the dendrogram exhibits the similarities of distance between soil samples or groups of soil samples. At the height of 3, the soil samples are grouped into 5 clusters. The number of clusters is reduced to 2 when the height approaches 6 and at the height of 8, only one cluster exists. In accordance with soil groups by the PCA biplot, two soil groups are clustered by the dendrogram. The soil groups will be characterized in the next section.

5.2.4 Soil groups by multivariate statistic tools

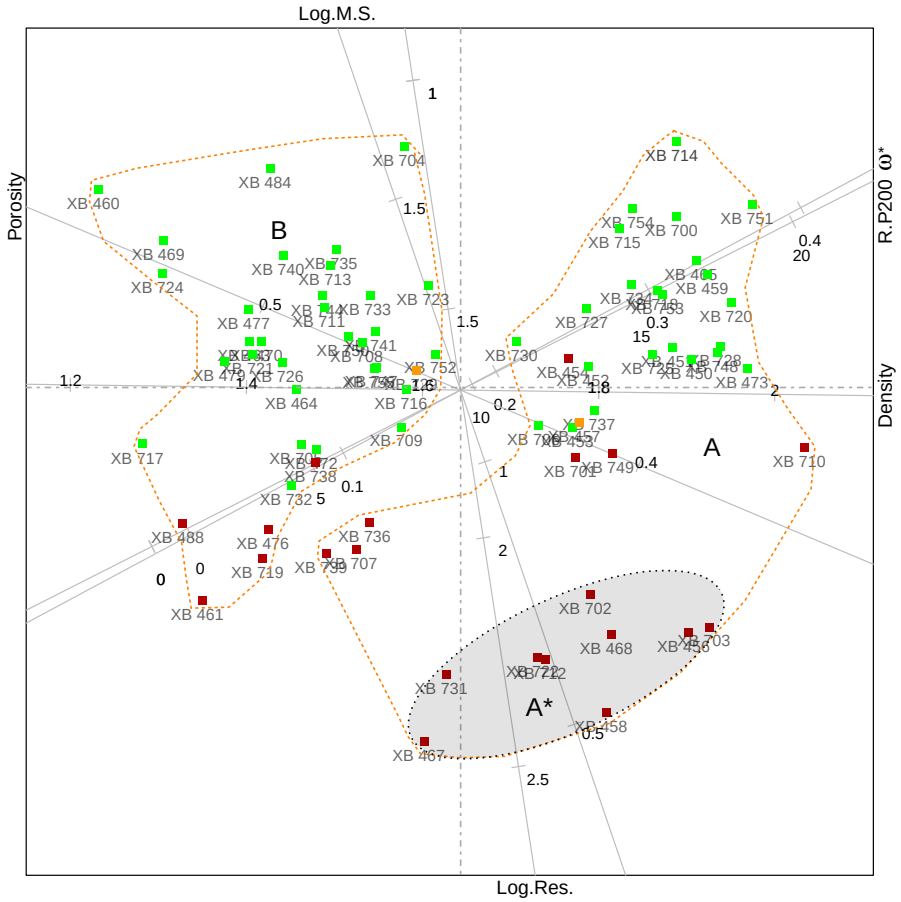


Figure 5.16: Clustered soil groups in a PCA biplot of the first two principal components.

In combination with the dendrogram, the soil samples can be clustered into groups in PCA biplot as shown in Figure 5.16. The second principal component, which pointing in vertical direction, enables a rough classification of the type of soils. Almost all soil samples that appear above the abscissa axis are classified as clay in green color and those below the abscissa are sand in red color. Two silt samples in orange color are located around the abscissa. However, some samples, which are located below close to the abscissa, are clay soils.

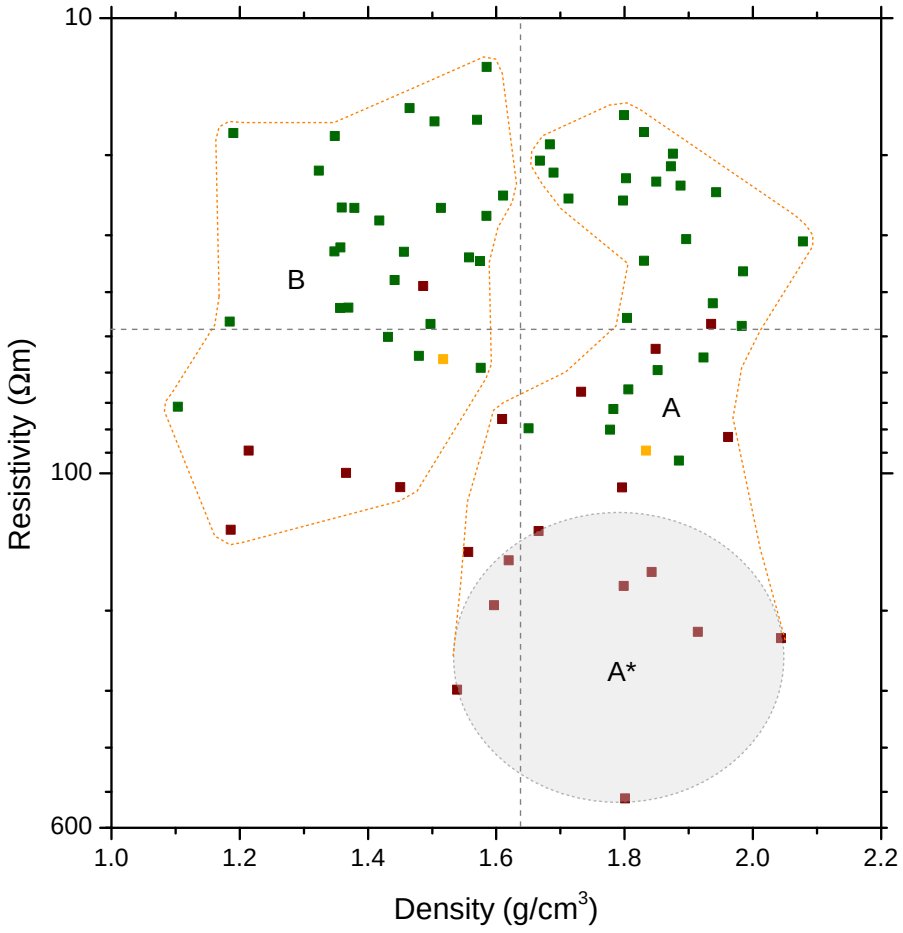


Figure 5.17: Soil groups in the cross-plot of density and resistivity.

The positions of the samples in the vertical direction are determined by the values of resistivity and magnetic susceptibility. This observation agrees with the

well-known trend that the resistivity increases as the grain size changes from clayey to sandy fraction. In the horizontal direction, where density has a greater influence, and to a smaller extent of porosity, volumetric water content and real part of relative dielectric permittivity, the trend is related to the depth of the corresponding soil samples. The soils on the left side with higher density originate from greater depth. Those on the right side with higher porosity and lower density were collected at a shallow depth.

With the reference to the dendrogram, the soils are clustered into two groups A and B as shown in Figures 5.15 and 5.16. The soil group B contains the soils that were located at shallow depth with low volumetric water content. It is obvious that these soils are collected at the dike body in shallow depth above water table. The soil group A contains soils from a greater depth with higher volumetric water content. The soil samples in group A were collected below the water table except the three sandy sample XB 707, XB 736 and XB 739 which originate from shallow depth. In group A, there exist a smaller soil group A* that contains sandy soils originating from a greater depth of 4 m as indicated in Figures 5.15 and 5.16.

Soil groups are clustered by six properties of soil samples in the PCA biplot. Figure 5.17 shows the cross-plot of density and resistivity in log scale of the soil samples. The groups of soils can be clustered in a similar way to the PCA biplot.

Chapter 6

Conclusions and Recommendations

In order to solve engineering and environmental problems, the application of geophysical techniques has been on a rapid rise in recent years. These non-invasive and cost effective techniques can be used to predict useful petrophysical and geotechnical engineering properties of soils in the subsurface. A successful application of geophysical methods in investigation of subsoil requires an adequate knowledge on petrophysical and geotechnical properties of the soil and their relationships.

Various types of soils were collected in different locations and depths from Vietnam and Germany. The soil samples originate from river dikes and adjacent foundation of civil engineering works.

Geotechnical properties of soil such as water content, density, Atterberg limits, clay fraction and shear strength parameters and hydraulic conductivity were investigated in laboratory. This study is focused on investigations on petrophysical and geotechnical properties of soils and their possible relations. Though, the hydraulic conductivity is an important parameter in dike inspection, this parameter was only determined for a small number of soil samples. Because of this limited number of

data, this parameter was not integrated in the multivariate statistical investigation. Considering the close relation between clay content and hydraulic conductivity, the parameter of clay content can be used to evaluate the hydraulic conductivity of the dike material.

Clay mineralogical analysis was performed on typical soil samples from Germany and Vietnam. The soils from Vietnam have higher clay fraction than those from Germany. The soils from Germany have kaolinite as dominant, while illite is present abundantly in the soils from Vietnam.

The results show that an increase of water content results in a decrease of shear strength parameters. Clay fraction presents a linear relation to liquid limit, plasticity index and specific surface area. Specific surface area shows a linear increase with clay fraction and liquid limit. These relations are in agreement with those known from literature. The soils from the dike body are clay with low to very low permeability. The logarithm of hydraulic conductivity of soils indicates a linear decrease with increasing clay fraction.

Petrophysical properties such as complex resistivity, dielectric permittivity, grain density, magnetic susceptibility and specific surface area were determined in the laboratory. The complex resistivity of soils was investigated in the frequency domain by spectral induced polarization (SIP). The phase shift between current and voltage signal of many soils from Vietnam are found to be negative in the low frequency range. CH_2-groups, alkyl chains or unsaturated hydrocarbon were found in these soils by organic chemical investigation. The negative phase may be a result of chemical reaction between clay minerals and organic matter under an applied current. Further investigations are need to find a plausible explanation for this phenomenon.

The soils from Vietnam with hight clay content have lower resistivity but higher dielectric permittivity compared to the soil samples from Germany. Especially, the soils from the two sites of An Bai and Dong Lam, Vietnam, have very high values of

imaginary part of dielectric permittivity that corresponds to a low electrical resistivity. The cross-plot of clay fraction and resistivity confirms the trend that an increase of clay fraction results in a decrease of resistivity. The volumetric water content has a strong effect on the real part of dielectric permittivity. Magnetic susceptibility and grain density of soils from Vietnam exhibit higher values than those of soils from Germany due to the higher clay content. Another reason may be the presence of iron oxide in the soils from Vietnam and the higher magnetic susceptibility of illite in comparison to kaolinite.

When dealing with a large amount of soil samples and more than three properties, statistical methods should be used to analyse and visualize the data and the relations among the properties in global view. Multivariate statistical methods of principal component analysis (PCA) and cluster analysis were applied to investigate the relation between geotechnical and petrophysical properties of soils. The soil samples and their properties are compiled in a data matrix, where the rows are soil samples and the their properties are compiled in the columns. Matrix manipulation algorithms are applied to reduce the dimensionality of the problem with the least loss of information. In statistical analysis of two variables, the coefficient of determination R^2 can be considered as goodness of fit of the model. In multivariate statistics, the new concepts of overall quality, adequacy and predictivity are used to access the goodness of fit of the model.

Both row and columns of the approximation matrix can be viewed as points and axes respectively in one graph as two- or three-dimensional biplot. The soil samples are presented as points while soil properties are the axes. A specific property of a soil sample can be interpolated by a projection of sample location onto the corresponding property axis. The cosine of angle between two axes exhibit the linear relation between those two properties.

Multivariate analysis methods offer potential tools to analyse and visualize a large

soil sample set with various properties. All soil samples and their properties can be visualized simultaneously in a global view. The technique also enables an effective classification of soil samples.

As demonstrated in a first example, the multivariate analysis of geotechnical data of soil samples from different locations in Vietnam has identified two groups of parameters. The correlation coefficient matrix indicates strong correlations between porosity, water content, density and friction angle. Most of these comprehensive correlations result from basic physical soil models. At the one hand, a larger porosity reduces the density and enables an increased volume of water in the pore space. On the other hand, increased water content reduces the friction angle. These four interrelated parameters show the strongest loadings in the first principal component. The other group of parameters, which dominate the second principal component, are cohesion and clay fraction that are characterized by a moderate correlation. It is known that increased clay content causes a rise in cohesion. Looking at the score plot of these two principal components, 93% of the total variance of all parameters and all samples is displayed. A cluster analysis based on the same geotechnical parameters provides a classification into eight soil clusters. The use of only one parameter of each group seems to be a less expensive alternative of soil classification. A cross-plot of the easily determinable parameters porosity and clay fraction enables a rough differentiation between sandy, silty and clayey soils.

A second example integrates petrophysical parameters that might be determined by geophysical field surveys along dikes or from geophysical logging in small boreholes. The geoelectrical method provides the resistivity of the soil material. The most relevant parameter that is extracted by a radar survey is the relative dielectric permittivity. The magnetic susceptibility is considered as additionally petrophysical parameter. The parameters porosity, natural raw density, and volumetric water content, which are determinable both from laboratory investigations or well logging,

complete this set of six parameters. In a similar way as in the first example, density, water content, and to less extent porosity dominate the first principal component. Because of the theoretically justified excellent correlation between relative dielectric permittivity and water content the real part of dielectric permittivity joins the first group of parameters. Resistivity and magnetic susceptibility, which show no correlation to the parameters of the first group, indicate the strongest loadings in the second principal component. The PCA biplot of the first two principal components provides 75% of the total variance of all samples. In combination of cluster analysis and PCA biplot, soils are classified as two clusters. Using a cross-plot of density and resistivity, the soil clusters are roughly identified.

The points representing soil samples on the biplot can be clustered into groups of soil. The different clusters reflect the type of soil and also the compaction or depth relative to the water table. However, the group boundaries are not fully consistent with the geotechnical soil classification.

The second example has demonstrated that the electrical resistivity is a key parameter in soil classification. This parameter is strongly related to clay content and water content of soils. Geoelectrical surveys along the crest of the dike, which are recommended for dike investigation, enable a sectioning into more sandy or clayey soils in the dike body.

A dike monitoring system, which is based on a permanent geoelectrical array at a fixed location, measures the changes in soil resistivity that can be converted into water content. The successful application of geoelectrical methods for dike inspection results in a wide acceptance of non-destructive geophysical tools for geotechnical problems.

Other geophysical methods that provide parameters like water content, clay content, porosity, and density should be integrated in inspection surveys to support a reliable soil classification.

References

AASHTO M145-91 (1991). *Classification of Soils and Soil-Aggregate Mixtures for Highway Construction Purposes.* American Association of State Highway and Transportation Officials, Washington, DC.

Al-Shayea, N. A. (2001). The combined effect of clay and moisture content on behavior of remolded unsaturated soils. *Engineering Geology*, 62:319–342.

ASTM D2216-98 (1998). *Standard Test Method for Laboratory Determination of Water (Moisture) Content of Soil and Rock by Mass.* American Society for Testing and Materials, West Conshohocken, PA.

ASTM D2487 (2000). *Standard Practice for Classification of Soils for Engineering Purposes (Unified Soil Classification System).* American Society for Testing and Materials, West Conshohocken, PA.

ASTM D422-63 (1998). *Standard Test Method for Particle-Size Analysis of Soils.* American Society for Testing and Materials, West Conshohocken, PA.

ASTM D4318-00 (2000). *Standard Test Method for Liquid Limit, Plastic Limit and Plasity Index of Soils.* American Society for Testing and Materials, West Conshohocken, PA.

Ayres, A. and Theilen, F. (2001). Natural gama-ray activity compared to geotechnical and environmental characteristics of nearsuface marine sediments. *Journal of Applied Geophysics*, 48:1–10.

Aysen, A. (2002). *Soil Mechanics - Basic Concepts and Engineering Applications*. A.A. Balkema Publishers.

Baecher, G. B. and Christian, J. T. (2003). *Realibility and Statistics in Geotechnical Engineering*. John Wiley & Sons Ltd, UK.

Balanis, C. A. (1989). *Advanced Engineering Electromagnetics*. John Willey & Sons, New York.

Banin, A. and Amiel, A. (1970). A correlative study of the chemical and physical properties of a group of natural soils of Israel. *Geoderma*, 3:185–197.

Baulieu, F. (1989). A classification of presence/absence dissimilarity co-efficients. *Journal of Classification*, 6:233–246.

Boadu, F. K. and Owusu-Nimo, F. (2010). Influence of petrophysical and geotechnical engineering properities on the electrical response of unconsolidated earth materials. *Geophyics*, 75(3):G21–G29.

Boadu, F. K. and Seabrook, B. C. (2006). Effect of clay content and salinity on the spectral electrical response of soils. *Journal of Environmental and Engineering Geophysics*, 11:161–170.

Börner, F. D. and Schön, J. H. (1991). A relation between the quadrature component of electrical conductivity and the specific surface area of sedimentary rocks. *The Log Analyst*, 32:612–613.

Braga, A., Malagutti, W., Dourado, J., and Chang, H. (1999). Correlation of electrical resistivity and induced polarization data with geotechnical survey stan-

dard penetration test measurements. *Journal of Environmental and Engineering Geophysics*, 4:123–130.

Brandes, I. M. (2005). *The negative chargeability of clays.* PhD thesis, The University of New South Wales, Australia.

Brown, B. and Mengel, J. (1983). Characteristics of red clay of douglas county, wisconsin. *Transportation Research Record*, 945:22–29.

BSI BS EN ISO14688-2 (2004). *Geotechnical investigation and testing. Identification and classification of soil. Principles for a classification.* British Standards Institution, London, United Kingdom.

Canan, B. (1999). *Dielectric properties of mixtures of clay-water-organic compounds.* PhD thesis, Dept. of Geophysics, Colorado School of Mines, Golden.

Canh, T., Bac, T. V., Tuyen, D. V., Weller, A., and Ngoc, P. Q. (2005). The combination of geophysical technologies to study instability of left Day river dike section in Y Yen, Nam Dinh province. *Journal of Geology*, series B 25:95–107.

Casagrande, A. (1948). Classification and identification of soils. *Trans. ASCE*, 113:901–991.

Cerato, A. and Lutenegger, A. (2005). Activity, relative activity and specific surface area of fine-grained soils. *Proceedings of the 16th International Conference on Soil Mechanics and Geotechnical Engineering (ICSMGE). Osaka, Japan*, 2:325–328.

Cosenza, P., Marmet, E., Reijiba, F., Cui, Y. J., Tabbagh, A., and Charlery, Y. (2006). Correlations between geophysical and geotechnical data: A case study at Garchy in France. *Journal of Applied Geophysics*, 60(3-4):165–178.

Das, B. M. (2007). *Fundamentals of Geotechnical Engineering.* The 3rd ed. CL-Engineering.

Das, B. M. (2008). *Advanced Soil Mechanics*. Third edition, Taylor & Francis, New York.

Davis, J. C. (2002). *Statistics and Data Analysis in Geology*. Third Edition, Johh Willey & Sons, New York.

Dearing, J. A., Hay, K. L., Baban, S. M. J., Huddleston, A. S., Wellington, E. M. H., and Loveland, P. J. (1996). Magnetic susceptibility of soil: An evaluation of conflicting theories using a national data set. *Geophysical Journal International*, 127:728–734.

Diep, N. V., Can, N. H., and Lai, H. V. (2004). European Flocods project - The Red River Delta, Vietnam in flood season. Technical report, Institute of Mechanics, Vietnamese Acadamy of Science and Technology.

DIN 18196 (2006). *Erd- und Grundbau - Bodenklassifikation für bautechnische Zwecke*. Deutsches Institut für Normung e.V, Beuth Verlag GmbH, Berlin.

Do, N. V. (1996). *Hydrogeological mapping in Nam Dinh and Thai Binh provinces*. Mineral Resources of Vietnam, Hanoi.

Dolinar, B. and Trauner, L. (2004). Liquid limit and specific surface of clay particles. *Geotechnical Testing Journal*, 27:1–5.

Dusseault, M. B. and Scafe, D. (1979). Mineralogical and engineering index properties of the Basal McMurray formation clay shales. *Canadian Geotechnical Journal*, 16:185–294.

Evans, M. E. and Heller, F. (2003). *Environmental Magnetism: Principles and Application of Enviromagnetics*. Elsevier Science (USA).

Farrar, D. M. and Coleman, J. D. (1967). The correlation of surface area with other properties of nineteen british clay soils. *Journal of Soil Science*, 18:118–124.

Fauchard, C. and Mériaux, P. (2007). *Geophysical and geotechnical methods for diagnosing flood protection dikes – Guide for implementation and interpretation.* Éditions Quae.

Gardner-Lubbe, S., le Roux, N. J., and Gower, J. C. (2008). Measures of fit in principal component and canonical variate analyses. *Journal of Applied Statistics*, 35:947–965.

Giao, P. H., Chung, S. G., Kim, D. Y., and Tanaka, H. (2003). Electric imaging and laboratory resistivity testing for geotechnical investigation of Pusan clay deposits. *Journal of Applied Geophysics*, 52(4):157–175.

Gordon, A. (1999). *Classification, 2nd ed.* Cliapnian and Hall/ CRC Press.

Gordon, A. D. (1987). A review of hierarchical classification. *Journal of the Royal Statistical Society*, 150:119–137.

GOST 25100-82 (1982). *Soil Classifications.* Russian State Standards, Moscow, Russia.

Gower, J. (1985). Measures of similarity, dissimilarity and distance. In Kotz, S., Johnson, N.L. and Read, C.B. (eds.). *Encyclopedia of Statistical Sciences*, 5:397–405.

Gower, J. C., Groenen, P. J. F., and Van der Velden, M. (2010). Area biplots. *Journal of Computational and Graphical Statistics*, 19(1):46–61.

Grim, R. E. (1962). *Applied Clay Mineralogy.* McGraw-Hill, New York.

Guéguen, Y. and Palciauskas, V. (1994). *Introduction to the Physics of Rocks.* Princeton University Press, Princeton.

Hair, J. F. and Anderson, R. E. (2010). *Multivariate Data Analysis*. 7th ed., Prentice Hall.

Hamilton, A. B. (1966). Freezing shrinkage in compacted clays. *Canadian Geotechnical Journal*, 3:1–17.

Holtz, R. D. and Kovacs, W. D. (1981). *An Introduction to Geotechnical Engineering*. Prentice-Hall Inc, New Jersey.

Jackson, D., Somers, K., and Harvey, H. (1989). Similarity coefficients: measures of co-occurrence and association or simply measures of occurrence. *The American Naturalist*, 133:436–453.

Johansson, S. and Dahlin, T. (1996). Seapage monitoring in an earth embankment dam by repeated resistivity measurement. *European Journal of Enviromental and Engeneering Geophysics*, 1:229–247.

Jolliffe, I. T. (2002). *Principal Component Analysis*. 2nd Ed. Springer-Verlag, New York, Inc.

Jones, D. P. (1997). Investigation of clay-organic reactions using complex resistivity. Master's thesis, Dept. of Geophysics, Colorado School of Mines, Golden.

Kato, S., Kawai, K., and Yoshimura, Y. (2000). *Estimation of Relation between Suction and Cohesion Using Unconfined Compression Tests*, pages 509–519. Unsaturated Soils for Asia. Balkema, Rotterdam.

Lambe, T. W. and Whitman, R. V. (1969). *Soil Mechanics*. John Willey & Sons, New York.

Lutenegger, A. J. and Cerato, A. B. (2001). Surface area and engineering properties of fine-grained soils. *Proceedings of the Fifteenth International Conference on Soil Mechanics and Geotechnical Engineering, Istanbul, Turkey*, 1:603–606.

Meunier, A. (2005). *Clays*. Springer-Verlag Berlin Heidenberg.

Mitchell, J. K. (1993). *Fundamentals of Soil Behavior*. Second edition, John Willey & Sons, Inc., New York.

Mitchell, J. K. and Soga, K. (2005). *Fundamentals of Soil Behavior*. Third edition, John Wiley & Sons, Inc., New York.

Moore, D. M. and Reynolds, R. C. (1997). *X-Ray Diffraction and the Identification and Analysis of Clay Minerals*. Oxford University Press, New York.

Morin, R. E. and Jacobs, H. S. (1984). Surface area determination of soils by adsorption of Ethylene Glycol Vapor. *Soil Science Society of America Proceedings*, 28:190–194.

Mortland, M. (1974). Specific surface and its relationships to some physical and chemical properties of soil. *Soil Science*, 78:343–348.

Mullins, C. E. (1977). Magnetic susceptibility of soils and its significance in soil science - a review. *Journal of Soil Science*, 28:223–246.

Murray, H. H. (2007). *Applied Clay Mineralogy: Occurrences, Processing and Application of Kaolins, Bentonites, Paligorskite-Sepiolite, and Common Clays* . Elsevier, Amsterdam, Netherlands.

Ngoc, P. Q. (2005). A geotechnical-geophysical study on the dike system in the North of Vietnam. Master thesis, Asian Institute of Technology, Bangkok, Thailand.

Niederleithinger, E., Weller, A., Lewis, R., Stoetzner, U., Fetchner, T., Lorenz, B., and Niessen, J. (2008). Evaluation of geophysical methods for river embankment investigation. *Proceeding of 4th International Symposium of Flood Defence. Toronto, Canada*.

Odell, R., Thornburn, T., and McKenzie, L. (1960). Relationships of Atterberg limits to some other properties of Illinois soils. *Soil Science Society of America Proceedings*, 24:297–300.

Olhoeft, G. R. (1985). Low-frequency electrical properties. *Geophysics*, 50(12):2492 – 2503.

Olphen, H. V. (1977). *An Introduction to Clay Colloid Chemistry*. Second edition, John Willey & Sons, New York.

Pansu, M. and Gautheyrou, J. (2006). *Handbook of Soil Analysis: Mineralogical, Organic and Inorganic Methods*. Springer- Verlag Berlin Heidelberg, New York.

Paquet, H. and Clauer, N. (1997). *Soils and Sediments: Mineralogy and Geochemistry*. Springer-Verlag Berlin Heidelberg.

Potter, D. K., Corbett, P. W. M., Barclay, S. A., and Haszeldine, S. (2004). Quantification of illite content in sedimentary rock using magnetic susceptibility - a rapid completement or alternative to x-ray difraction, 74 (5), p.730-735. *Journal of Sedimentary Research*, 74 (5):730–735.

Rabitti, S., Boldrin, A., and Vitturi, L. (1983). Relationships between surface area and grain size in bottom sediments. *Journal of Sedimentary Research*, 53:665–667.

Revil, A. and Glover, P. W. J. (1997). Theory of ionic-surface electrical conduction in porous media. *Physical Review B*, 55:1757–1773.

Reyment, R. A. and Savazzi, E. (1999). *Aspects of Multivariate Statistical Analysis in Geology*. Elsevier Science B.V, Amsterdam.

Saarenketo, T. (1998). Electrical properties of water in clay and silty soils. *Journal of Applied Geophysics*, 40:73–88.

Salat, C. and Junge, A. (2010). Dielectric permittivity of fine-grained fractions of soil samples from eastern Spain at 200 MHz. *Geophysics*, 75:J1–J9.

Schön, J. H. (1996). *Physical Properities of Rocks - Fundamentals and Principles of Petrophysics*, volume 18. Elsevier Science Ltd.

Schwartz, B. F., Schreiber, M. E., and Yan, T. (2008). Qualifying field-scale soil moisture using electrical resistivity imaging. *Journal of Hydrology*, 362:234–246.

Shaw, D. J. (1992). *Introduction to Colloid and Surface Chemistry*. Fourth edition, Butterworth-Heinemann Ltd., Oxford.

Sjödahl, P., Dahlin, T., S.Johansson, and Loke, M. (2008). Resistivity monitoring for leakage and internal erosion detection at Haellby embankment dam. *Journal of Applied Geophysics*, 65:155–164.

Smith, C., Hadas, A., Dan, J., and Koyumdjisky, H. (1985). Shrinkage and Atterberg limits relation to other properties of principle soil types in Israel. *Geoderma*, 35:47–65.

Sneath, P. H. A. and Soka, R. R. (1973). *Numerical Taxonomy*. W.H. Freeman and Co., San Francisco.

Sridharan, A. and Prakash, K. (1998). Mechanism controlling the shrinkage limit of soils. *Geotechnical Testing Journal*, 21:240–250.

Sudha, K., Israil, M., Mittal, S., and Rai, J. (2008). Soil charaterization using electrical resistivity tomography and geotechnical investigations. *Journal of Applied Geophysics*, 67:74–79.

Swan, A. R. H. and Sandilands, M. (1995). *Introduction to Geological Data Analysis*. Blackwell Science Ltd.

Takahashi, T. and Yamamoto, T. (2010). An attempt at soil profiling on a river embankment using geophysical data. *Exploration Geophysics*, 41:102–108.

TCVN 4198:1995 (1995). *Soils - Laboratory methods of determination of grain size distribution*. TCVN , Vietnamese Directorate for Standards and Quality, Hanoi, Vietnam.

Terzaghi, K., Peck, R. B., and Mesri, G. (1996). *Soil Mechanics in Engineering Practice*. Third Edition, Johh Willey & Sons, New York.

Thermann, K., Gau, C., and Tiedemann, J. (2009). *Shear strength parameters from direct shear tests - Influencing factors and their significance*. In: Engineering Geology Special Publications. The Geological Society of London, UK.

Tinh, D. Q. (2001). Participatory planning and management for flood mitigation and prepareness and trends in the Red River Basin, Vietnam. Technical report, Department of Dike Management and Flood Control, Hanoi, Vietnam.

Topp, G. C., Davis, J. L., and Annan, A. P. (1980). Electromagnetic determination of soil water content: Measurements in coaxial transmission lines. *Water Resources Research*, 16:574 –582.

Tuyen, D. V., Canh, T., and Weller, A. (2000). Geophysical investigation of river dikes in Vietnam. *European Journal of Environmental and Engineering Geophysics*, 4:195–206.

Van Dam, R., Hendrickx, J., Harrison, B., Borchers, B., Norman, D., Ndur, S., Jasper, C., Niemeyer, P., Nartey, R., Vega, D., Calvo, L., and Simms, J. (2004). Spatial variability of magnetic soil properties. In *Detection and remediation technologies for mines and minelike targets VIII, Proceedings of the SPIE*.

Velde, B. and Meunier, A. (2008). *The Origin of Clay Minerals in Soils and Weathered Rocks*. Springer-Verlag Berlin Heidelberg.

Wackernagel, H. (2003). *Multivariate Geostatistics - An Introduction with Applications*. Third Edition, Springer-Verlag Berlin Heidelberg, Berlin.

Warkentin, B. (1968). Properties of clay and silt soils of the Eastern Abitibi area in Quebec. *Soil Res. Bulletin. Department of Soil Science. Macdonald College of McGill University Quebec*, page 20.

Warkentin, B. (1972). Use of liquid limit in characterizing clay soils. *Canadian Journal of Soil Science*, 52:457–464.

Weller, A. and Börner, F. (1996). Measurements of induced polarisation for environmental purposes. *Environmental Geology*, 27:329–334.

Weller, A., Canh, T., Breede, K., and Vu, N. T. (2006). Multi-electrode measurements at Thai Binh dikes (Vietnam). *Near Surface Geophysics*, 4:135–143.

Weller, A., Canh, T., Toan, D. V., Tuyen, D. V., Bac, T. V., Ngan, P. K., and Dung, N. Q. (1996). The effect of combination of geophysical methods to find out defects in the dike body in Hanoi, Vietnam. *Journal of Earth Science*, 18(4):339–348.

Weller, A., Lewis, R., and Niederleithinger, E. (2008). *GeophysikalishceVerfahren zur Struturerkundung und Schwachstellenanalyse von Flussdeichen – ein Handbuch*. Bundesanstalt für Materialforschung und -prüfung (BAM), Berlin.

Weller, A., Möller, M., Lewis, R., and Canh, T. (2010). Dike monitoring at Red River by geophysical and geotechnical tools. In *Near Surface 2010 – 16th European Meeting of Environmental and Engineering Geophysics, Zurich, Switzerland*.

Yin, J.-H. (1999). Properties and behaviour of Hong Kong marine deposits with different clay contents. *Canadian Geotechnical Journal*, 36:1085–1095.

Yong, R. N. and Warkentin, B. P. (1975). *Soil Properties and Behaviour*. Elsevier Scientific Publishing Company, Amsterdam.

Appendix

Descriptive summary of geotechnical properties of soil samples from Vietnam

Parameters	CF	LL	PL	PI	Wc	Porosity	Density	Coh.	F.Ang
	(%)	(%)	(%)	(%)	(%)	(%)	(g/cm^3)	(kPa)	(Deg)
					Ngo Xa (n=15)				
Mean	20.06	48.17	27.67	20.51	37.00	50.66	1.80	17.36	14.38
Standard Deviation	5.60	4.90	3.28	2.18	6.08	3.55	0.07	9.69	2.72
Standard Error	1.40	1.26	0.85	0.56	1.57	0.92	0.02	2.50	0.70
Min	9.00	41.20	22.70	15.00	26.10	44.10	1.69	6.87	10.43
First Quartile	16.00	44.20	25.30	19.25	32.85	48.25	1.76	10.10	11.99
Median	18.50	48.40	28.00	20.50	37.40	50.90	1.78	11.77	14.78
Third Quartile	25.00	51.20	29.55	22.00	42.20	53.50	1.86	25.95	16.65
Max	29.00	57.50	33.80	23.70	46.00	56.20	1.91	38.36	18.98
					An Bai (n=51)				
Mean	20.29	39.22	23.64	15.58	36.13	50.21	1.82	8.61	10.99
Standard Deviation	10.07	4.91	2.71	2.91	7.97	5.74	0.10	5.10	5.29
Standard Error	1.41	0.74	0.41	0.44	1.12	0.80	0.01	0.71	0.74
Min	6.00	27.20	18.40	8.80	22.70	39.10	1.70	2.52	5.60
First Quartile	11.00	36.68	20.90	13.15	28.85	46.10	1.73	5.30	6.67
Median	17.00	38.55	24.15	15.95	38.90	53.20	1.76	6.93	7.74
Third Quartile	30.50	44.35	25.73	18.23	43.35	55.00	1.88	9.86	16.09
Max	36.00	46.90	28.60	20.30	46.50	57.20	2.03	20.96	20.25
					Yen Dinh (n=9)				
Mean	11.89	39.08	23.65	15.43	34.24	49.13	1.83	3.88	13.94
Standard Deviation	8.46	0.51	0.26	0.25	7.19	5.09	0.08	1.92	7.21
Standard Error	2.82	0.25	0.13	0.13	2.40	1.70	0.03	0.64	2.40
Min	4.00	38.50	23.40	15.10	27.50	44.20	1.72	1.86	5.48
First Quartile	4.00	38.80	23.48	15.33	28.40	44.60	1.74	2.26	6.73
Median	8.00	39.05	23.60	15.45	29.40	45.90	1.88	2.84	19.14
Third Quartile	18.00	39.33	23.78	15.55	41.60	54.40	1.89	5.59	19.85
Max	24.00	39.70	24.00	15.70	44.30	55.90	1.90	6.38	20.98
					Tra Linh (n=10)				
Mean	25.60	42.83	26.59	16.21	38.28	51.53	1.80	7.77	10.30
Standard Deviation	7.28	7.52	4.07	3.87	10.10	5.93	0.10	3.20	4.33
Standard Error	2.30	2.38	1.29	1.22	3.19	1.88	0.03	1.01	1.37
Min	11.00	32.70	22.50	10.00	25.90	43.10	1.61	3.63	4.39
First Quartile	23.25	36.83	23.33	13.88	31.40	48.03	1.76	5.47	7.49
Median	25.00	42.00	25.35	16.35	36.75	51.20	1.81	6.72	9.33
Third Quartile	30.00	47.65	29.25	18.33	39.43	53.20	1.84	10.33	13.99
Max	37.00	56.60	33.00	23.70	58.50	62.40	1.94	12.56	17.49
					Vu Doai (n=12)				
Mean	15.92	35.07	21.30	13.77	31.20	47.73	1.85	6.79	13.52
Standard Deviation	6.82	1.85	1.72	1.28	5.41	3.36	0.05	4.61	5.52
Standard Error	1.97	0.70	0.65	0.48	1.56	0.97	0.01	1.33	1.59
Min	4.00	31.30	18.90	12.30	21.00	42.50	1.76	0.39	5.10
First Quartile	14.25	34.70	20.20	12.95	28.45	44.65	1.81	2.50	9.72
Median	18.50	35.90	21.00	13.90	30.30	48.35	1.85	7.16	13.70
Third Quartile	19.50	36.10	22.60	14.10	33.60	49.38	1.89	10.18	17.55
Max	23.00	36.70	23.60	16.10	42.50	54.30	1.92	14.52	21.60
					Dong Lam (n=11)				
Mean	15.30	37.00	22.34	14.66	35.78	50.86	1.79	5.49	12.25
Standard Deviation	9.91	4.62	1.45	3.20	6.32	4.87	0.08	2.33	8.25
Standard Error	2.99	1.63	0.51	1.13	1.91	1.47	0.02	0.70	2.49
Min	1.00	26.00	19.00	7.00	26.80	43.60	1.72	1.28	5.37
First Quartile	5.00	37.43	22.30	15.13	30.00	46.45	1.73	3.43	6.50
Median	21.00	38.60	22.65	15.85	39.50	54.00	1.75	6.57	6.50
Third Quartile	23.00	39.28	23.13	16.10	41.00	54.60	1.86	6.97	20.85
Max	26.00	40.20	23.50	16.70	42.70	55.00	1.92	8.34	24.51

Note: n: Number of samples, CF: Clay fraction, LL: Liquid Limit, PL: Plastic Limit, PI: Plastic Index, Wc: Water content, Coh: Cohesion, and F.Ang: Angle of Internal Friction